



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁵ : C07K 13/00, C12N 15/31, 15/62 A61K 37/02, C07K 3/18	A1	(11) International Publication Number: WO 93/22342 (43) International Publication Date: 11 November 1993 (11.11.93)
(21) International Application Number: PCT/SE93/00375 (22) International Filing Date: 28 April 1993 (28.04.93) (30) Priority data: 9201331-7 28 April 1992 (28.04.92) SE (71) Applicant (for all designated States except US): HIGHTECH RECEPTOR AB [SE/SE]; c/o Active, Skeppsbron 2, S- 211 20 Malmö (SE). (72) Inventors; and (75) Inventors/Applicants (for US only) : BJÖRCK, Lars [SE/ SE]; Kornvägen 40, S-240 17 Södra Sandby (SE). SJÖ- BRING, Ulf [SE/SE]; Lilla Sigridsgatan 1, S-223 50 Lund (SE). (74) Agents: BERG, S., A. et al.; H. Albihns Patentbyrå AB, Box 3137, S-103 62 Stockholm (SE).		(81) Designated States: JP, US, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE). Published <i>With international search report.</i>
(54) Title: PROTEIN L AND HYBRID PROTEINS THEREOF (57) Abstract The invention relates to sequences of protein L which bind to light chains of immunoglobulins. The invention also relates to hybrid proteins thereof which are able to bind to both light and heavy chains of immunoglobulin G, in particular protein LG. The invention also relates to DNA-sequences which code for the proteins, vectors which include such DNA-sequences, host cells which have been transformed with the vectors, methods for producing the proteins, reagent appliances for separation and identification of immunoglobulins, compositions and pharmaceutical compositions which contain the proteins.		

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	FR	France	MR	Mauritania
AU	Australia	GA	Gabon	MW	Malawi
BB	Barbados	GB	United Kingdom	NL	Netherlands
BE	Belgium	GN	Guinea	NO	Norway
BF	Burkina Faso	GR	Greece	NZ	New Zealand
BG	Bulgaria	HU	Hungary	PL	Poland
BJ	Benin	IE	Ireland	PT	Portugal
BR	Brazil	IT	Italy	RO	Romania
CA	Canada	JP	Japan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo	KR	Republic of Korea	SE	Sweden
CH	Switzerland	KZ	Kazakhstan	SK	Slovak Republic
CI	Côte d'Ivoire	LJ	Liechtenstein	SN	Senegal
CM	Cameroon	LK	Sri Lanka	SU	Soviet Union
CS	Czechoslovakia	LU	Luxembourg	TD	Chad
CZ	Czech Republic	MC	Monaco	TG	Togo
DE	Germany	MG	Madagascar	UA	Ukraine
DK	Denmark	ML	Mali	US	United States of America
ES	Spain	MN	Mongolia	VN	Viet Nam
FI	Finland				

Protein L and Hybrid Proteins Thereof

5 The present invention relates to sequences of protein L which bind to light chains of immunoglobulins. The invention also relates to hybrid proteins of protein L having the ability to bind to light chains of all Ig and also to bind to light and heavy chains of immunoglobulin G, DNA-sequences which code for the proteins vectors that contain such DNA-sequences, host cells transformed
10 by the vectors, methods for preparing the proteins, reagent apparatus for separating and identifying immunoglobulins, compositions and pharmaceutical compositions which contain the proteins.

15 The invention relates in particular to the DNA-sequence and to the amino acid sequence of the light-chain forming domains of protein L.

20 Proteins which bind to the constant domains (of high affinity) of the immunoglobulins (Ig) are known. Thus, protein A (from Staphylococcus aureus) (Forsgren, A. and Sjöquist, J. (1966) Protein A from Staphylococcus aureus. I. Pseudo-immune reaction with human gamma-globulin. J. Immunol. 97: 822-827) binds to IgG from
25 various mammal species. The binding of protein A to IgG is mediated essentially via surfaces in the Fc-fragment of the heavy chain of the IgG-molecule, although a certain bond is also effected with surfaces in the Fab-fragment of the IgG. Protein A lacks the ability of
30 binding to human IgG3 and neither will it bind to IgG from several other animal species, such as important laboratory animals, for instance rats and goats, which limits the use of protein A.

35 Protein G (Björck, L. and Kronvall, G. (1984) Purification and some properties of streptococcal protein G, a

SUBSTITUTE SHEET

novel IgG-binding reagent. J. Immunol. 133: 969-974;
Reis, K., Ayoub, E. and Boyle, M. (1984) Streptococcal
Fc receptors. I. Isolation and partial characterization
of the receptor from a group C streptococcus. J.

5 Immunol. 132: 3091-3097) binds to heavy chains in human
IgG and to all four of its subclasses and also to IgG
from most mammals, including rats and goats.

10 Protein H (Åkesson, P., Cooney, J., Kishimoto, F. and
Björck, L. (1990) Protein H - a novel IgG binding bacte-
rial protein. Molec. Immun. 27: 523-531) binds to the
Fc-fragment in IgG from human beings, monkeys and rab-
bits. However, the bond is weaker than in the case of
protein G and A, which may be beneficial when wishing to
15 break the bond with a weak agent, for instance when
purifying proteins which are readily denatured with the
aid of antibodies.

20 Protein M (Applicant's Patent Application PCT/SE
91100447) binds to the Fc-fragment in IgG from humans,
monkeys, rabbits, goats, mice and pigs.

25 Protein L (Björck, L. (1988) Protein L, a novel bacteri-
al cell wall protein with affinity to Ig L chains. J.
Immunol. 140: 1194-1197), which binds to the light
chains in immunoglobulins from all of the classes G, A,
M, D and E is known (USP 4,876,194). The amino acid se-
quence and the binding domains of this protein, however,
have hitherto been unknown.

30

The aforesaid proteins can be used in the analysis,
purification and preparation of antibodies and for
diagnostic and biological research.

35 The elimination of immunoglobulins, with the aid of

SUBSTITUTE SHEET

plasmapheresis, can have a favourable effect on some autoimmune diseases. A broadly binding protein would be an advantage when wishing to eliminate all classes of antibodies in this context.

5

It has long been known that infectious conditions can be prevented or cured with the introduction of an immune serum, i.e. a serum which is rich in antibodies against the organism concerned or its potentially harmful product. Examples hereof are epidemic jaundice, tetanus, diphtheria, rabies and generalized shingles. Antibodies against a toxic product may also be effective in the case of non-infectious occasioned conditions. Serum produced in animals against different snake venoms is the most common application in this respect. However, the administration of sera or antibody preparations is not totally without risk. Serious immunological reactions can occur in some cases. Singular cases of the transmission of contagious diseases, such as HIV and hepatitis through the agency of these products have also been described. In order to avoid these secondary effects, it has been desirable to produce therapeutic antibodies in test tubes. A large number of novel techniques for the preparation of antibodies in test tubes have been proposed in recent years. Examples of such techniques are hybridom techniques, synthesis of chima-antibodies and the preparation of antibodies in bacteria. These techniques also enable antibodies to be specially designed which can further widen the use of such molecules as therapeutics, for instance in the case of certain tumour-diseases. In the case of some of these novel methods, however, the product totally lacks the Fc-fragment to which all of the described IgG-binding proteins, with the exception of protein L, bind. There is consequently a need of a process for purifying anti-

10

15

20

25

30

35

SUBSTITUTE SHEET

bodies for therapeutic use, wherein proteins which have a broad binding activity/specificity, can be of value.

5 It has long been possible to utilize the antibody reaction with its high grade specificity for diagnosing past or, in some cases, ongoing infections with different parasites. This indirect method of indicating infectious agents is called serology and, in many cases, may be the only diagnostic alternative. In certain cases, it can
10 also be of interest to exhibit specific IgE- or IgA-antibodies. When diagnosing with the aid of serology, the antigen is most often fastened to a solid phase, whereafter serum taken from the patient is incubated with the antigen. Antibodies that have been bound from
15 the patient can then be detected in different ways, often with the aid of a secondary antibody (for instance, an antibody which is directed against the light chains of human antibodies) to which an identifiable label has been attached, such as alkaline phosphatase,
20 biotin, radioactive isotopes, fluorescein, etc. In this context, a protein having a broad Ig binding capacity can be used as an alternative to secondary antibodies.

25 There are a number of non-therapeutic and non-diagnostic reasons for the necessity to bind antibodies. Antibodies are often used in research, both for detection and for purifying the antigen against which they are directed. All techniques which facilitate the purification of antibodies and, in particular, techniques which enable
30 different classes to be purified, are of interest in this context.

Consequently, there is a serious need of a protein which has a broad binding activity/specificity and which binds
35 to several different classes of immunoglobulins from different animal species. At present, there is no known

SUBSTITUTE SHEET

protein which will bind to all immunoglobulin classes. The earlier known proteins A, G, H and M bind only to heavy chains in IgG. The known protein L (Björck et al, 1988) binds to the light χ -chains and γ -chains in immunoglobulins of all classes, although the bonds are much weaker on the κ -chains. Applicant has charted protein L, has determined the amino acid sequence for protein L, has identified the light-chain binding domains on protein L, and has used these to produce hybrid proteins which possess the IgG-Fc-binding domains of protein G. The Applicant is able to show through protein LG that a protein of broader binding activity/ specificity can be produced thereby. The aforesaid proteins A, G, H and M bind to the same surfaces, or to very closely lying surfaces on IgG-Fc. The protein L which binds to light chains can thus be combined with any other functionally similar protein which binds to the Fc-fragment of heavy chains. A similar broadening of the Ig-binding activity is achieved with all alternatives.

Thus, the present invention relates to the sequence of protein L which binds to light chains in Ig and has the amino acid sequence disclosed in Figure 1, and variants, subfragments, multiples or mixtures of the domains B1-B5 having the same binding properties. The invention also relates to a DNA-sequence which codes for such protein sequences, for instance the DNA-sequence in Figure 1.

The invention is concerned with a hybrid protein which is characterized by comprising domains which bind to the light χ -chains and λ -chains in immunoglobulins of all classes, and also comprises domains which bind to heavy chains in immunoglobulin G, wherein those domains which bind to the light chains are chosen from among the B1-, B2-, B3-, B4- and B5-domains in protein L (see Claim 1)

SUBSTITUTE SHEET

and those domains which bind to heavy chains of immunoglobulins are chosen from the C1-, C2- and C3-domains in protein G; the A-, B- and C1-domains from protein H; the A-, B1-, B2- and S-domains in protein M1 or the E-, D-, A-, B- and C-domains in protein A (see Figure 6) and variants, subfragments, multiples or mixtures of these domains that have the same binding properties which bind to heavy chains of immunoglobulins.

By subfragment is meant a part-fragment of the given domains or fragments which include parts from the various domains having mutually the same binding properties. By variants is meant proteins or peptides in which the original amino acid sequence has been modified or changed by insertion, addition, substitution, inversion or exclusion of one or more amino acids, although while retaining or improving the binding properties. The invention also relates to those proteins which contain several arrays (multiples) of the binding domains or mixtures of the binding domains with retained binding properties. The invention also relates to mixtures of the various domains of amino acid sequences having mutually the same binding properties.

The invention relates in particular to a hybrid protein designated LG, and is characterized in that the hybrid protein includes the B-domains in protein L which bind to the light chains in immunoglobulins, and the C1-domains and C2-domains in protein G which bind to heavy chains and have the amino acid sequence disclosed in Figure 3. The invention also relates to variants, subfragments, multiples or mixtures of these domains.

Protein LG is a hybrid protein having a molecular weight of about 50 kDa (432 amino acids) and comprising four domains, each of which binds to light chains in immuno-

globulins, and two IgG-binding domains from protein G. The hybrid protein combines a broad IgG-binding activity, deriving from the high-grade binding ability of protein G to the Fc-fragment of the heavy chain on IgG with the ability of the protein L to bind to light chains of all classes of immunoglobulins. Thus, protein LG binds polyclonal human IgG, IgM, IgA, IgD and IgE. The affinity for human polyclonal IgG is $2 \times 10^{10} \text{ M}^{-1}$. All four human immunoglobulin classes are bound. Binding to human IgG is effected with both the κ - and the λ -chain. Both the Fc-fragment and the Fab-fragment of IgG are bound to the hybrid protein. The protein also binds human IgA-, IgD-, IgE- and IgM-antibodies. The bond is stronger to human immunoglobulins which carry χ than to those which carry the λ -isotope of light chains. IgG from most mammals will be bound by protein LG, thus also IgG from goats and cows, which do not bind to protein L. However, rabbit-IgG which binds relatively weakly to protein L will bind well to the fusion protein. IgM and IgA-antibodies from mice, rats and rabbits will be bound to the protein.

Protein LG is highly soluble. It is able to withstand heat and will retain its binding properties even at high temperatures. The binding properties also remain in a broad pH-range of 3-10. The protein withstands detergent and binds marked or labelled proteins subsequent to separation in SDS-PAGE and transference to membranes with elektroblotting. The protein can be immobilized on a solid phase (nitrocellulose, Immobilon®, polyacrylamide, plastic, metal and paper) without losing its binding capacity. The binding properties are not influenced by marking with radioactive substances, biotin or alkaline phosphatase. (The binding abilities of the protein LG are disclosed in Example 3).

SUBSTITUTE SHEET

The protein comprises 432 amino acids and has a molecular weight of 50 kDa deriving therefrom. The sequence is constructed of an ala sequence of the three last amino acids in the A-domain of the protein L (val-glu-asn),
5 this ala sequence being unrelated to the two proteins, whereafter the four mutually high-grade homologous B-domains from protein L follow. The first of the B-domains is comprised of 76 amino acids, and the remaining domains are each comprised of 72 amino acids. The first
10 nine amino acids from the fifth B-domain are included and followed by two non-related amino acids (pro-met). The protein G-sequences then follow. The last amino acid in the so-called S-domain from protein G is followed by an IgG-binding domain from protein G (C1; 55 amino
15 acids), the intermediate D-region (15 amino acids) and the second IgG-binding C-domain (C2; 55 amino acids). The last amino acid is a methionine, which occurs in natural protein G as the first amino acid in the so-called W-region.

20

The invention also relates to DNA-sequences which code for the aforesaid proteins.

25

The gene which codes for the IgG-binding amino acid sequences can be isolated from the chromosomal DNA from Staphylococcus aureus based on the information on the DNA-sequence for protein A (S. Löfdahl, B. Guss, M. Uhlen, L. Philipsson and M. Lindberg. 1983. Gene for staphylococcal protein A. Proc. Natl. Acad. Sci. USA.
30 80: 697-701) and Figure 6, or from G-streptococcus, preferably strain G 148 or C-streptococcus, preferably strain Streptococcus equisimilis C 40, based on the information on protein G (B. Guss, M. Eliasson, A. Olsson, M. Uhlen, A.-K. Frej, H. Jörvall, I. Flock
35 and M. Lindberg. 1986. Structure of the IgG-binding

SUBSTITUTE SHEET

regions of streptococcal protein G. EMBO. J. 5: 1567-1575) and Figure 6, or from group A-streptococcus, e.g. S. pyogenes (type M1) based on the information on the DNA-sequence for protein H (H. Gomi, T. Hozumi, S. Hattori, C. Tagawa, F. Kishimoto and L. Björck. 1990. The gene sequence and some properties of protein H - a novel IgG binding protein J. Immunol. 144: 4046-4052) and Figure 6, or from the chromosomal DNA in group A-streptococcus type M1 based on the information on the DNA-sequence for protein M (Applicant's Patent Application, PCT/SE 91100447) and Figures 6 and 7. The gene which codes for the protein that binds to light chains can be isolated from the chromosomal DNA from Pepto-coccus magnus 312 based on the information on the DNA-sequence for protein L in Claim 2.

By using the chromosomal DNA't obtained from the aforesaid bacteria as a template, a DNA-fragment defined with the aid of two synthetic oligonucleotides can then be specifically amplified with the aid of PCR (Polymerase Chain Reaction). This method also enables recognition sites to be incorporated for restriction enzymes in the ends of the amplified fragments (PCR technology, Ed: PCR Technology. Principles and Applications for DNA Amplification. Ed. Henry Erlich. Stockton Press, New York, 1989). The choice of recognition sequences can be adapted in accordance with the vector chosen to express the fragment or the DNA-fragment or other DNA-fragments with which the amplified fragment is intended to be combined. The amplified fragment is then cleaved with the restriction enzyme or enzymes concerned and is combined with the fragment/the other fragments concerned and the fragments are then cloned together in the chosen vector (in this case, the expression vector) (Sambrook, J.E. Fritsch and T. Maniatis, 1989, Molecular cloning: A laboratory manual, 2nd Ed. Cold Spring Harbor Laborato-

SUBSTITUTE SHEET

- ries, Cold Spring Harbor, New York, USA). The plasmid vector pHD313 can be used (Dalbøge, H.E. Bech Jensen, H. Töttrup, A. Grubb, M. Abrahamson, I. Olafsson and S. Carlsen, 1989. High-level expression of active human cystatin C in *Escherichia coli*. *Gene*, 79: 325-332), alternatively one of the vectors in the so-called PET-series (PET 20, 21, 22, 23) retailed by Novagen (Madison, Wisconsin, USA).
- 10 The hybrid proteins are then incorporated in an appropriate host, preferably *E. coli*. The invention also relates to such hosts as those in which the hybrid proteins are incorporated.
- 15 Those clones which produce the desired proteins can be selected from the resultant transformants with the aid of a known method (Fahnestock et al., *J. Bacteriol.* 167, 870 (1986)).
- 20 When the proteins that can bind to the light chains in the immunoglobulins and to the heavy chains in IgG have been purified from the resultant positive clones with the aid of conventional methods, the binding specificities of the proteins are determined for selection of those clones which produce a protein that will bind to the light chains in immunoglobulins and to the heavy chains in IgG.
- 25
- 30 Subsequent to having isolated plasmid DNA't in said clone with conventional methods, the DNA-sequence in the inserted material is determined with known methods (Sanger et al., *Proc. Natl. Acad. Sci. USA* 74, 5463 (1977)).
- 35 The invention also relates to DNA-sequences which hybridize with said identified DNA-sequences under conven-

tional conditions and which code for a protein that possesses the desired binding properties. Strict hybridizing conditions are preferred.

5 Expression of the genes can be effected with expression vectors which have the requisite expression control regions, the structural gene being introduced after said regions. As illustrated in Figure 1 and Claim 2, the structural gene can be used for protein LG or other
10 hybrid proteins with protein L.

With regard to expression vectors, different host-vector-systems have been developed, of which the most suitable host-vector-systems can be selected for expres-
15 sion of the genes according to the present invention.

The present invention also relates to a method of producing the inventive hybrid proteins by cultivating a host cell which is transformed with an expression vector
20 in which DNA't which codes for the proteins according to the invention is inserted.

This method includes the steps of

25 (1) inserting into a vector a DNA-fragment which codes for the hybrid proteins;

(2) transforming the resultant vector into an appropriate host cell;

30 (3) cultivating the resultant, transformed cell for preparation of the desired hybrid protein; and

(4) extracting the protein from the culture.

35

SUBSTITUTE SHEET

In the first step, the DNA-fragment which codes for the hybrid protein is inserted in a vector which is suitable for the host that is to be used to express the hybrid protein. The gene can be inserted by cleaving the vector with an appropriate restriction enzyme, and then ligating the gene with the vector.

In the second step, the vector with the hybrid plasmid is inserted into host cells. The host cells may be Escherichia, coli, Bacillus subtilis or Saccharomyces cerevisiae or other suitable cells. Transformation of the expressions hybrid vector into the host cell can be effected in a conventional manner and clones which have been transformed can then be selected.

In the third step, the obtained transformants are cultivated in an appropriate medium for preparation of the desired proteins by expression of the gene coded for the hybrid protein.

In the fourth step, the desired protein is extracted from the culture and then purified. This can be achieved with the aid of known methods. For instance, the cells can be lysed with the aid of known methods, by treating the cells with ultrasonic sound, enzymes or by mechanical degradation. The protein which is released from the cells or which excretes in the medium can be recovered and purified with the aid of conventional methods often applied within the biochemical field, such as ion-exchange chromatography, gel filtration, affinity chromatography with the use of immunoglobulins as ligands, hydrophobic chromatography or reverse-phase chromatography. These methods can be applied individually or in suitable combinations.

SUBSTITUTE SHEET

As before mentioned, the inventive proteins may be used for binding, identifying or purifying immunoglobulins. They can also be bound to pharmaceuticals and used in formulations which have delayed release properties. To
5 this end, the protein may be present in a reagent appliance for pharmaceutical composition in combination with appropriate reagents, additives or carriers.

The proteins can be handled in a freeze-dried state or
10 in a PBS-solution (phosphate-buffered physiological salt solution) pH 7.2 with 0.02% NaN_3 . It can also be used connected to a solid phase, such as carbohydrate-based phases, for instance CNBr-activated sepharose, agarose, plastic surfaces, polyacrylamide, nylon, paper, magnetic
15 spheres, filter, films. The proteins may be marked with biotin, alkaline phosphatase, radioactive isotopes, fluorescein and other fluorescent substances, gold particles, ferritin, and substances which enable luminescence to be measured.

20 Other proteins may also be used as carriers. These carriers may be bound to or incorporated in the proteins, in accordance with the invention. For instance, it is conceivable to consider the whole of proteins A,
25 G, H, M as carriers for inserted sequences of protein L which bind to light chains. In turn, these carriers can be bound to the aforesaid carriers.

The pharmaceutical additions that can be used are those
30 which are normally used within this field, such as pharmaceutical qualities of mannitol, lactose, starch, magnesium stearate, sodium saccharate, talcum, cellulose, glyucose, gelatine, saccharose, magnesium carbonate and similar extenders, such as lactose, dicalcium phosphate and the like; bursting substances, such as starch
35 or derivatives thereof; lubricants such as magnesium

SUBSTITUTE SHEET

stearate and the like; binders, such as starch, gum arabicum, polyvinylpyrrolidone, gelatine, cellulose and derivatives thereof, and the like.

5 The invention will now be described in more detail with reference to the accompany drawings, in which

Figure 1 illustrates the plasmid pHD389; the ribosomal binding sequence, the sequence for the signal peptide from ompA and recognition sequence for several restriction enzymes are shown;

Figure 2 illustrates the amino acid and nucleic acid sequence for protein LG;

15 Figure 3 is a schematic overall view of the production of protein L;

Figure 4 is a schematic overall view of the production of protein LG;

Figures 5a, 5b and 5c are schematic overall views of the production of the hybrid proteins LA, LM and LH respectively;

25 Figure 6 is a schematic inclusive illustration of protein A, G, H and M1. IgGFc-binding domains are for protein A: E, D, A, B and C; for protein G: C1, C2 and C3; for protein H: A and/or B; and for protein M1: A, B1, B2, B3 and S;

30 Figure 7 illustrates the amino acid and nucleic acid sequence for protein M1;

SUBSTITUTE SHEET

Figure 8 illustrates Western Blot for protein G, L and LG with certain immunoglobulins and immunoglobulin fragments; and

5 Figure 9 illustrates Slot-Blot for protein L, G and LG with IgG, IgX and Ig Fc.

The amino acid and nucleic acid sequence of the light-chain binding domains of protein L is illustrated in
10 Claims 1 and 2 respectively.

It will be observed that the drawings are not to scale.

Example 1

15

Cloning and expression of the IgG-light-chain-binding domains in Protein L

20 **Construction of synthetic oligonucleotides (primers) for amplifying sequences coded for protein L, domain B1-B4**

It has been found that a protein L peptide (expressed in E. coli) constructed of the sequence ala-val-glu-asn-domain B1 (from protein L) binds to the light chains of the immunoglobulins (W. Kastern, U. Sjöbring and L. Björck. 1992. Structure of peptostreptococcal protein L and identification of a repeated immunoglobulin light chain-binding domain. J. Biol. Chem. in-print). Since
25 this simple protein L-domain has a relatively low affinity to Ig, ($1 \times 10^7 \text{ M}^{-1}$), and since the naturally occurring protein L which is constructed of several mutually
30 similar domains (B1-B5) has a high affinity to Ig ($1 \times 10^{10} \text{ M}^{-1}$) four of these domains have been expressed together in the following way:

35

PL-N and PL-C1 are synthetic oligonucleotides (manufactured by the Biomolecular Unit at Lund University (Sweden) in accordance with Applicant's instructions) which have been used to amplify a clonable gene fragment which is amplified with PCR (Polymerase Chain Reaction) and which codes for four Ig-binding protein L domains (ala-val-glu-asn-B1-B2-B3-B4-lys-lys-val-asp-glu-lys-pro-glu-glu). Amino acids in the protein L-sequence are given for the primer which corresponds to the coded strand (PL-N):

PL-N: 5'-GCTCAGGCGGCGCCGGTAGAAAATAAAGAAGAAACACCAGAAAC-3'
valgluasnlysglugluthrproglu

5'-end of this oligonucleotide is homologous with the coded strand in the protein L-gene (emphasized): those codons which code for the last three amino acids in the A-domain (val-glu-asn) are followed by the codons for the first six amino acids in the first of the Ig-binding domains in protein L (B1).

PL-C1: 5'-CAGCAGCAGGATTCTTATTATTCTTCTGGTTTTTCGTCAACTTTCTT-3'

This oligonucleotide is homologous with the opposing non-coding strand in the gene for protein L (the sequence corresponds to the first nine amino acids in domain B5).

DNA-fragments which have been amplified with the aid of PL-N contain the recognition sequence for the restriction enzyme HpaII (emphasized) immediately before the codon which is considered to code for the first amino acid (val) in the expressed protein L-fragment. The fragment which is cleaved with HpaII can be ligated with

SUBSTITUTE SHEET

5 DNA (in this case, consisting of the used expression vector pHD389) which has been cleaved with the restriction enzyme NarI. The DNA-fragment that has been cleaved with HpaII and ligated with vector pHD389, which has been cleaved with NarI, will be translated in the correct reading frame. The construction results in translation of an additional amino acid (ala) immediately in front of the first amino acid in protein L.

10 DNA-fragments which have been amplified with the aid of PL-C1 will contain the recognition sequence for the restriction enzyme BamHI (overlined above the sequence) immediately after the sequence which codes for the last amino acid in the expressed protein L-fragment (glu).
15 The vector pHD389 contains a unique recognition sequence for BamHI as part of its so-called multiple cloning sequence which follows the NarI recognition sequence. DNA-fragments which have been amplified with the aid of PL-C1 will include two so-called stop-codons (empha-
20 sized) which results in translation of the fragment inserted in the vector to cease.

The sequence which was considered to be amplified contains no internal recognition sequences for the restriction enzymes HpaII or BamHI.
25

Amplifying and cloning procedures

(PCR) (Polymerase Chain Reaction) was effected with a
30 protocol described by Saiki, R.D. Gelfand, S. Stoffel, S. Scharf, R. Higuchi, G. Horn, K. Mullis and H. Erlich, 1988; Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science 239: 487-49127; PCR was effected in a Hybaid Intelligent Heating-
35 block (Teddington, UK): 100 μ l of a reaction mixture contained 50 mM KCl, 10 mM Tris-HCl, pH 8.3, 1.5 mM

SUBSTITUTE SHEET

MgCl₂, 100 µ/ml gelatine, 300 µM with respect to each of the deoxynucleotides (dATP, dCTP, dGTP, dTTP), (Pharmacia), 20 pmol of each of the oligonucleotides PL-N and PL-C1, 10 µl of a target (template) DNA-solution containing 0.1 mg/ml of chromosomal DNA from Peptostreptococcus magnus, species 312. The mixture was covered with mineral oil (Sigma) and DNA't was denatured by heating to 98°C for 10 minutes. 2.5 units of AmpliTaq (Perkin Elmer Cetus, Norwalk, CT) were added and PCR was then carried out with 25 cycles consisting of a denaturing step at 94°C for 1 minute, followed by a hybridizing step at 56°C for 1 minute, and finally by an extension step at 72°C for 1 minute. Amplified DNA was analyzed by electrophoresis in agarose gel. The amplified DNA't was cleaved with the restriction enzymes HpaII (Promega), (8 units/µg amplified DNA) and BamHI (Promega), (10 units/µg amplified DNA) at 37°C. The thus amplified and subsequently cleaved DNA-product was isolated by electrophoresis in a 2% (weight by volume) agarose gel (NuSieve agarose, FMC Biproducts) in a TAE-buffer (40 Mm Tris, 20 Mm Na-acetate, 2 Mm EDTA, Ph 8.0). The resulting 930 base-pair large fragment was cut from the gel. The DNA concentration in this removed gel-piece was estimated to be 0.05 mg/ml. The agarose-piece containing the cleaved, amplified fragment was melted in a water bath at 65°C, whereafter the fragment was allowed to cool to 37°C. 10 µl (0.5 µg) of this DNA was transferred to a semimicrotube (Sarstedt) preheated to 37°C, whereafter 1 µl of the vector pHD389 was immediately added and cleaved with NarI (Promega) and BamHI, 1 µl 10xligas-buffer (Promega) and 1 µl T4 DNA-ligase (Promega; 1 unit/µl). The ligating reaction was then used to transform E. coli, strain LE392, which had been competent in accordance with the rubidium/calcium-chloride-method as described by Kushner (1978). Molecular biological standard methods have been used in the manipulation

SUBSTITUTE SHEET

of DNA (Sambrook, J.E. Fritsch and T. Maniatis, 1989. Molecular cloning: A laboratory manual. 2nd Ed. Cold Spring Harbor Laboratories, Cold Spring Harbor, New York, USA). The cleaving and ligating conditions recommended by the manufacturer of DNA-ligase and restriction enzymes have been followed in other respects.

Expression system

10 The vector pHD389 (see Figure 2) is a modified variant of the plasmid pHD313 (Dalbøge, H.E. Bech Jensen, H. Töttrup, A. Grubb, M. Abrahamson, I. Olafsson and S. Carlsen, 1989. High-level expression of active human cystatin C in *Escherichia coli*. *Gene*, 79: 325-332). The
15 vector, which is replicated in *E. coli* (contains ori = origin of replication from plasmid pUC19) is constructed so that DNA-fragments which have been cloned into the cleaving site of *NarI* will be transcribed and translated downstream of and in the immediate vicinity of the
20 signal peptide (21 amino acids), from envelope-protein *ompA* from *E. coli*. Translation will be initiated from the codon ATG which codes for the first amino acid (methionine) in the signal peptide. This construction permits the translated peptide to be transported to the
25 periplasmic space in *E. coli*. This is advantageous, since it reduces the risk of degradation of the desired product of enzymes occurring intracellularly in *E. coli*. Moreover, it is easier to purify peptides which have been exported to the periplasmic space. Unique recognition sequences (multiple cloning sequences) for several other restriction enzymes, among them *ecoRI*, *SalI* and *BamHI* are found immediately after the *NarI* cleaving
30 site. An optimized so-called Shine-Dalgarno-sequence (also called ribosomal binding site, RBS) is found seven
35 nucleotides upstream from the ATG-codon in the signal sequence from *ompA*, this optimized sequence binding to a

SUBSTITUTE SHEET

complementary sequence in 16S rRNA in the ribosomes and is responsible for the translation being initiated in the correct place. The transcription of such DNA as that which is co-transcribed with the signal sequence for ompA is controlled by the P_R -promotor from coliphage λ . The vector also contained the gene for cI857 from coliphage λ whose product down-regulates transcription from P_R (and whose product is expressed constitutively). This cI857-mediated down-regulation of transcription from P_R is heat-sensitive. The transcription regulated from this promotor is terminated with the aid of a so-called rho-independent transcription terminating sequence (forms a structure in DNA't which results in the DNA-dependent RNA-polymerase leaving the DNA-strand) which is placed in the vector immediately downstream of the multiple cloning sequence. The plasmid also carries the β -lactamase gene (from the plasmid pUC19) whose product permits ampicillin-selection of E. coli clones that have been transformed by the vector.

20

Selection of protein L-producing clones

The transformed bacteria are cultivated, or cultured, on culture plates with an LB-medium which also contained ampicillin in a concentration of 100 μ g/ml. Cultivation of the bacteria progressed overnight at 30°C, whereafter the bacteria were transferred to an incubator where they were cultivated for a further 4 hours at 42°C. The plates were kept in a refrigerator overnight. On the next day, the colonies were transferred to nitrocellulose filters. Filters and culture plates were marked so as to enable the transferred colonies to be readily identified on respective culture plates. The culture plates were again incubated overnight at 30°C, so that remaining rests of transferred bacteria colonies could again grow. The plates were then kept in a refrigerator.

SUBSTITUTE SHEET

The bacteria in the colonies on the nitrocellulose-
impressions were lysed by incubating the filter in 10%
SDS for 10 minutes. Filters containing lysed bacteria
were then rinsed with a blocking buffer which comprised
5 PBS (pH 7.2) with 0.25% gelatine and 0.25% Tween-20
(four baths, 250 ml each at 37°C), whereafter the filter
was incubated with radioactively marked (marked with
¹²⁵I in accordance with the chloramin-T-method) Ig-κ-
chains (20 ng/ml in PBS with 0.1% gelatine). The incuba-
10 tion took place at room temperature over a period of 3
hours, whereafter non-bound radioactively marked protein
was rinsed-off with PBS (pH 7.2) containing
0.5 M NaCl, 0.25% gelatine and 0.25% Tween-20 (four
baths, 250 ml each at room temperature). All filters
15 were exposed to X-ray film. Positive colonies were
identified on the original culture plate. Clones which
reacted with Ig-κ-chains were selected and analyzed with
respect to the size on the DNA-fragment introduced in
the vector. One of these clones was selected for the
20 production of protein L, pHDL. The DNA't introduced from
this clone into plasmid pHD389 was sequenced. The DNA-
sequence was found to be in full agreement with corre-
sponding sequences (B1-B4 and 21 bases in B5) in the
gene for protein L from Peptostreptococcus magnus,
25 strain 312. The size and binding properties of the
protein produced by clone pHDL was analyzed with the aid
of SDS-PAGE (see Figure 8), dot-blot experiment (see
Figure 9) and competitive binding experiments.

30 Production of protein L

Several colonies from a culture plate with E. coli pHDL
were used to inoculate a preculture (LB-medium with an
addition of 100 mg/l ampicillin), which was cultured at
35 28°C overnight. On the following morning, the preculture
was transferred to a larger volume (100 times the volume

of the preculture) of fresh LB-medium containing ampicillin (100 mg/l) and was cultured in shake-flasks (200 rpm), (or fermentors) at 28°C. The culture temperature was raised to 40°C (induction of transcription) when the absorbency value at 620 nm reached 0.5. Cultivation then continued for 4 hours (applied solely to cultivation in shake-flasks). Upon completion of the cultivation process, the bacteria were centrifuged down. The bacteria were then lysed with an osmotic shock method at 4°C (Dalbøge et al., 1989 supra). The lysate was adjusted to a pH = 7. Remaining bacteria rests were then centrifuged down, whereafter the supernatant was purified on IgG-sepharose in accordance with earlier described protocol for protein G and protein L (U. Sjöbring, L. Björck and W. Kastern. 1991. Streptococcal protein G: Gene structure and protein binding properties. J. Biol. Chem. 266: 399-405; W. Kastern, U. Sjöbring and L. Björck. 1992. Structure of peptostreptococcal protein L and identification of a repeated immunoglobulin light chain-binding domain. J. Biol. Chem. in-print.

The expression system gave about 20 mg/l of protein L when cultivation in shake-flasks. The culture was deposited at DSSM, Identification Reference DSSM E. coli LE392/pHDL.

SUBSTITUTE SHEET

Example 2

Cloning and expression of protein LG

5 **Construction of oligonucleotides (primers) for amplifying sequences which code for protein LG**

Protein L

10 It has been found that a protein L-peptide (expressed in
15 E. coli) constructed of the sequence ala-val-glu-asn-
domain B1 (from protein L) will bind to the light chains
of the immunoglobulins (Kastern, Sjöbring and Björck,
1992, J. Biol. Chem. in-print). Since the affinity of
this simple domain to Ig is relatively low ($1 \times 10^{-7} \text{ M}^{-1}$)
and since the naturally occurring protein L, which is
comprised of several mutually similar domains (B1-B5)
has a higher affinity to Ig ($1 \times 10^{10} \text{ M}^{-1}$), four of
these domains have been expressed together in the fol-
20 lowing way:

25 PL-N and PL-C2 are synthetic oligonucleotides (manufactured at the Biomolecular Unit at Lund University (Sweden) in accordance with Applicant's instructions) which were used, with the aid of PCR (Polymerase Chain Reaction) to amplify a clonable gene fragment, called B1-4, which codes for four Ig-binding protein L domains (ala-val-glu-asn-B1-B2-B3-B4-lys-lys-val-asp-glu-lys-pro-glu-glu):

30
PL-N: 5'-GCTCAGGCGGCGCCGGTAGAAAATAAAGAAGAAACACCAGAAAC-3'
valgluasnlysglugluthrproglu

P1-C2: 5'-CAGCAGCAGCCATGGGTTCTTCTGGTTTTTCGTCAACTTTCTTA-
35 3'

Amino acids have been shown under corresponding triplets in the coded strand. DNA-fragments which have been amplified with the aid of PL-N contain the recognition sequence for the restriction enzyme *HpaII* immediately upstream of the triplet which codes for the first amino acid (val) in the expressed protein L-fragment. The fragment that has been cleaved with *HpaII* can be ligated with DNA (in this case, the used expression vector pHD389) which has been cleaved with *NarI*. The construction results in translation of an extra amino acid (ala) immediately upstream of the first amino acid in the protein L-fragment. The DNA-fragment that has been amplified with the aid of PL-C2 will contain the recognition sequence for the restriction enzyme *NcoI* (emphasized) immediately downstream of the sequence which codes for the last amino acid in the expressed protein L-fragment (glu). Amplified fragments which have been cleaved with *NcoI* can be ligated to the *NcoI*-cleaved, PCR-generated protein-asp-CDC-met-fragment (see below).

20

Protein G

It is known that a simple C-domain from protein G will bind to IgG (B. Guss, M. Eliasson, A. Olsson, M. Uhlen, A.-K. Frej, H. Jörnvall, I. Flock and M. Lindberg. 1986. Structure of the IgG-binding regions of streptococcal protein G. EMBO. J. 5: 1567-1575). The strength at which a simple C-domain binds to IgG is relatively low ($5 \times 10^7 \text{ M}^{-1}$). A fragment which consists of two C-domains with an intermediate D-region having a length of 15 amino acids, however, has a considerably higher affinity to IgG ($1 \times 10^9 \text{ M}^{-1}$). CDC-N and CDC-C are oligonucleotides which have been used as PCR-primers to amplify a clonable DNA-fragment, designated CDC, which

35

SUBSTITUTE SHEET

codes for two IgG-binding protein G-domains (pro-met-asp-CDC-met).

CDC-N: GGCCATGGACACTTACAAATTAATCCTTAATGGT
metaspthrtyrlysleuileleuasngly

CDC-C: CAGGTCGACTTATTACATTTTCAGTTACCGTAAAGGTCTTAGT

10 Amino acids in the resultant sequence have been shown
beneath the primer of the coding strand. DNA-fragments
which have been amplified with the aid of CDC-N contain
the recognition sequence for the restriction enzyme NcoI
(marked with a line above the sequence). Cleaved ampli-
fied fragments can be ligated with the fragment that has
15 been amplified with the aid of PL-C2 and then cleaved
with NcoI. The fragment will therewith be translated to
the correct reading frame. DNA-fragments which have been
amplified with the aid of CDC-C will contain two so-
called stop condons (emphasized) which terminate trans-
20 lation. The recognition sequence for the restriction
enzyme SalI (marked with a line above the sequence)
follows immediately afterwards, this sequence also being
found in the expression vector pHD389 (see Figure 1).

25 Those sequences which code for the binding properties of
protein L (B1-B5) and for protein G (CDC) respectively
contain no internal recognition sequences for the re-
striction enzymes HpaII, SalI or NcoI.

30 Amplification and cloning procedures

PCR (Polymerase Chain Reaction) was carried out in
accordance with a protocol described by Saiki et al.,
1988; PCR was carried out in a Hybaid Intelligent Heat-
35 ing-block (Teddington, UK): 100 µl of the reaction
mixture contained 50 mM KCl, 10 mM Tris-HCl, pH 8.3, 1.5

5 mM $MgCl_2$, 100 $\mu g/ml$ gelatine, 300 μM with respect to each of the deoxynucleotides (dATP, dCTP, dGTP, dTTP), (Pharmacia). In order to amplify sequences which code for the light-chain binding parts of protein L, there were added 20 pmol of each of the oligonucleotides PL-N and PL-C2, and 10 μl of a DNA-solution which contained 0.1 mg/ml of chromosomal DNA from Peptostreptococcus magnus, strain 312. By way of an alternative, 20 pmol were added to each of the oligonucleotide pairs CDC-N and CDC-C and 10 μl of a DNA-solution which contained 0.1 mg/ml of chromosomal DNA from a group C streptococcus strain (Streptococcus equisimilis) called C40 (U. Sjöbring, L. Björck and W. Kästern. 1991. Streptococcal protein G: Gene structure and protein binding properties. J. Biol. Chem. 266: 399-405 or with NcoI and SalI (10 U/ μg PCR-product), (for CDC) at 37°C. The thus amplified and subsequently cleaved DNA-fragments were then separated by electrophoresis in a 2% (weight by volume) agarose gel (NuSieve agarose, FMC Bioproducts) in a TAE-buffer (40 mM Tris, 20 mM acetic acid, 2 mM EDTA, pH 8.0). The resultant fragments, 930 bp (for B1-4) and 390 bp (for CDC) were cut from the gel. The concentration of DNA in the thus separated gel pieces was estimated to be 0.05 mg/ml. The agarose pieces cut from the gel and containing the cleaved, amplified fragments (B1-4 and CDC) were melted in a water bath at 65°C, whereafter they were allowed to cool to 37°C. 10 μl (0.5 μg) of this DNA were transferred to a semimicrotube (Sarstedt), preheated to 37°C, whereafter 1 μl of the vector pHD389 which had been cleaved with NarI and SalI were added. 1 μl 10 x ligase buffer (Promega) and 1 μl T4 DNA-ligase (1 unit/ μl) were also added. The ligating reaction was permitted to take place at 37°C for 6 hours. The cleaving and ligating conditions recommended by the producer of DNA-ligase and restriction enzymes (Promega) were followed in other respects. The

SUBSTITUTE SHEET

ligating reaction was then used to transform E. coli, strain LE392, which had been made competent in accordance with the rubidium-chloride/calcium-dichloride method as described by Kushner (1978). Manipulation of DNA was effected in accordance with molecular biological standard methods (Sambrook et al., 1989).

Expression system

The vector pHD389 (see Figure 2) is a modified variant of the plasmid pHD313 (Dalbøge et al., 1989). The vector which was replicated in E. coli (contains origin of replication from plasmid pUC19) is constructed such that DNA-fragments which have been cloned in the cleaving site for NarI will be expressed immediately after, or downstream, of the signal peptide (21 amino acids) from the envelope protein ompA from E. coli. Translation will be initiated from the ATG-codon which codes for the first amino acid (methionine) in the signal peptide. The construction with an E. coli-individual signal sequence which precedes the desired peptide enables the translated peptide to be transported to the periplasmic space in E. coli. This is beneficial since it reduces the risk of degradation of the desired product through the intracellular occurrent enzymes of E. coli. Furthermore, it is easier to purify peptides which have been exported to the periplasmic space. Unique recognition sequences (multiple cloning sequences) for several other restriction enzymes, among them EcoRI, SalI and BamHI are present immediately downstream of the NarI cleaving site. An optimized so-called Shine-Dalgarno sequence (also called ribosomal binding site, RBS) is found seven nucleotides upstream of the ATG-codon in the signal sequence from ompA, this optimized Shine-Dalgarno sequence binding to a complementary sequence in 16S rRNA in the ribosomes and in a manner to decide that the

SUBSTITUTE SHEET

translation is initiated in the correct place. The transcription of such DNA as that which is co-transcribed with the signal sequence for ompA is controlled by the P_R -promotor from coliphage λ . The vector also contains the gene for cI857 from coliphage λ , the product of which regulates-down transcription from P_R and the product of which is expressed constitutively. This cI857-mediated down-regulation of transcription from P_R is heat-sensitive. Transcription which is regulated, or controlled, from this promotor will be terminated with the aid of a so-called rho-independent transcription terminating sequence which is inserted in the vector immediately downstream of the multiple cloning site. The plasmid also carries the gene for β -lactamase (from the plasmid pUC19), the product of which permits ampicillin-selection of E. coli clones that have been transformed with the vector.

Selection of protein LG-produced clones

The transformed bacteria are cultivated on culture plates with LB-medium which also contained ampicillin in a concentration of 100 μ g/ml. The bacteria were cultivated overnight at 30°C, whereafter they were transferred to a cultivation cabinet (42°C) and cultured for a further four (4) hours. The plates were stored in a refrigerator overnight. On the following day, the colonies were transferred to nitrocellulose filters. The filters and culture plates were marked, so that the transferred colonies could later be identified on the culture plate. The culture plates were again incubated overnight at 30°C, so that rests of transferred bacteria colonies remaining on the plates could again grow. The plates were then stored in a refrigerator. The filter was incubated in 10% SDS for 10 minutes, so as to lyse the bacteria in the colonies on the nitrocellulose

impression. Filters containing lysed bacteria were then rinsed with a blocking buffer consisting of PBS (pH 7.2) with 0.25% gelatine and 0.25% Tween-20 (four baths of 250 ml at 37°C), whereafter the filter was incubated with radioactively (marked with ^{125}I according to the chloramine-T-method) marked Ig- κ -chains (20 ng/ml) in PBS with 0.1% gelatine). The incubation process took place at room temperature for four (4) hours, whereafter non-bound radioactively marked protein was rinsed-off with PBS (pH 7.2) containing 0.5 M NaCl, 0.25% gelatine and 0.25% Tween-20 (four baths, 250 ml each at room temperature). All filters were exposed to X-ray film. Positive colonies on the original culture plate were identified. A number of positive colonies were re-cultivated on new plates and new colony-blot experiments were carried out with these plates as a starting material with the intention of identifying E. coli colonies which bind IgG Fc. These tests were carried out in precisely the same manner as that described above with respect to the identification of E. coli-colonies which expressed Ig light-chain-binding protein, with the exception that a radioactively marked (^{125}I) IgG Fc (20 ng/ml) was used as a probe. Clones which reacted with both proteins were selected and analyzed with regard to the size of the DNA-fragment introduced in the vector. One of these clones was chosen for production of protein LG, pHDLG. The DNA't taken from this clone and introduced into plasmid pHD389 was sequenced. The DNA-sequence exhibited full agreement with corresponding sequences (B1-B4 and 21 bases in B5) in the gene for protein L from Peptostreptococcus magnus, strain 312, and with C1DC2 sequence in group C streptococcus strain C40. The size and binding properties of the protein produced from clone pHDLG was analyzed with the aid of SDS-PAGE (see Figure 8), dot-blot experiment (see Figure 10) and competitive binding experiments.

SUBSTITUTE SHEET

Production of protein LG

Several colonies from a culture plate with *E. coli* PHDLG were used to inoculate a preculture (LB-medium with an addition of 100 mg/l ampicillin) were cultivated at 28°C overnight. In the morning, the preculture was transferred to a larger volume (100 times the volume of the preculture) of fresh LB-medium containing ampicillin (100 mg/l) and was cultivated in vibrating flasks (200 rpm), (or fermenters) at 28°C. When an absorbance value of 0.5 was reached at 620 nm, the cultivation temperature was raised to 40°C (induction of transcription). The cultivation process was then continued for 4 hours (applies only to cultivation in vibrated flasks). The bacteria were centrifuged down upon termination of the cultivation process. The bacteria were then lysed at 4°C in accordance with an osmotic shock method (Dalbøge et al., 1989). The lysate was adjusted to a pH of 7. Remaining bacteria rests were centrifuged down and the supernatant then purified on IgG-sepharose, in accordance with the protocol earlier described with reference to protein G and protein L. (Sjöbring et al., 1991, Kastern et al., 1992).

The expression system gave about 30 mg/l of protein LG when cultivation in vibrated flasks. A deposition has been made at DSSM, Identification Reference DSSM *E. coli* LE392/PHDLG.

30 Example 3

Analysis of the binding properties of protein LG

Western Blot

35 Protein G (the C1DC2-fragment), protein L (four B-

SUBSTITUTE SHEET

domains) and protein LG were isolated with SDS-PAGE (10% acrylamide concentration). The isolated proteins were transferred to nitrocellulose membranes in three similar copies (triplicate). Each of these membranes was incubated with radioactively marked proteins (20 ng/ml: one of the membrane-copies was incubated with human polyclonal IgG, another with human IgG Fc-fragment and the third with isolated human IgG χ chains. Non-bound radioactively marked proteins were rinsed off and all filters were then exposed to X-ray film.

Slot-blot

Human polyclonal Ig-preparations and Ig-fragments were applied with the aid of a slot-blot appliances on nitrocellulose filters in given quantities (see Figure 10) on three similar copies. Each of these membranes was incubated with radioactively marked proteins (20 ng/ml). One of the membrane copies was incubated with protein LG, another with protein L and the third with protein G. Non-bound radioactively marked proteins were rinsed-off and all filters were then exposed to X-ray film.

The results are shown in Figures 9 and 10.

Other binding experiments have been carried out, with the following results:

SUBSTITUTE SHEET

TABLE

Binding of the proteins G, L and LG to immunoglobulins.

Binding protein:	G	K _a	L	K _a	LG	K _a
<u>Immunoglobulin</u>						
<u>Human:</u>						
Polyclonal IgG*	+	67 (10)	+	9.0	+	20
IgG subclasses	+	2.0	+		+	
IgG ₁	+	3.1	+		+	
IgG ₂	+	6.1	+		+	
IgG ₃	+	4.7	+		+	
IgG ₄	+					
IgG fragment						
Fc*	+	6.0 (0.5)	-			+
F(ab') ₂ *	+	0.4 (0.2)	+		+	
kappa	-		+	1.5	+	
lambda	-		(-) [#]			
Other Ig-classes						
IgM	-		+	11.6	+	
IgA	-		+	10.4	+	
IgE	-		+		+	
IgD	-					
<u>Other Species:</u>						
Polyclonal						
Monkey	+		+		+	
Rabbit						
IgG	+	70	+	0.074	+	
IgG-Fc	+	3.0	-		+	
IgG-F(ab') ₂	+	0.44			+	
Mouse	+	41	+	2.6	+	
Rat	+	1.5	+	0.39	+	
Goat	+	14	-		+	

SUBSTITUTE SHEET

TABLE (cont'd.)

Binding of the proteins G, L and LG to immunoglobulins.

Binding protein:		G	K _a	L	K _a	LG	K _a
Immunoglobulin							
Bovine	IgG ₁	+	3	-		+	
	IgG ₂	+	2	-		+	
Horse		+		-		+	
Guinea Pig		+		+		+	
Sheep		+		-		+	
Dog		+		-		+	
Pig		+		+		+	
Hamster		+					
Cat		-		-			
Hen		-		-			
Monclonals ^{&}							
Mouse							
	IgG ₁	+		+	1	+	
	IgG _{2a}	+		+		+	
	IgG _{2b}	+				+	
	IgG ₃	+				+	
	IgM	-		+		+	
	IgA	-		+		+	
Rat							
	IgG _{2a}	+		+		+	
	IgG _{2b}	+				+	
	IgG _{2c}	+				+	

SUBSTITUTE SHEET

K_a = affinity constant (M^{-1}). * The numerals within parenthesis disclose the affinity of a recombinant protein G comprised of two IgG-binding domains. # A weak bond to lambda chains exists. & Binding to P1 and PLG depends on the type of light chain of Ig.

It will thus be seen that the synthesized hybrid protein LG has a broad binding activity/specificity.

SUBSTITUTE SHEET

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:

- (A) NAME: HighTech Receptro AB
- (B) STREET: c/o Active, Skeppsbron 2
- (C) CITY: MALMO
- (E) COUNTRY: SWEDEN
- (F) POSTAL CODE (ZIP): 211 20
- (G) TELEPHONE: 040/35 07 00
- (H) TELEFAX: 040/ 23 74 05
- (I) TELEX: 32637 Active S

(ii) TITLE OF INVENTION: Hybridprotein

(iii) NUMBER OF SEQUENCES: 1

(iv) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Floppy disk
- (B) COMPUTER: IBM PC compatible
- (C) OPERATING SYSTEM: PC-DOS/MS-DOS
- (D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(EPO)

(v) CURRENT APPLICATION DATA:

APPLICATION NUMBER: SE PCT/SE93/00375

(vi) PRIOR APPLICATION DATA:

- (A) APPLICATION NUMBER: SE 9201331-7
- (B) FILING DATE: 28-APR-1992

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 305 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Escherichia coli LE392/pHDL, DSM 7054

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

Ala	Val	Glu	Asn	Lys	Glu	Glu	Thr	Pro	Glu	Thr	Pro	Glu	Thr	Asp	Ser
1				5				10						15	
Glu	Glu	Glu	Val	Thr	Ile	Lys	Ala	Asn	Leu	Ile	Phe	Ala	Asn	Gly	Ser
			20					25					30		
Thr	Gln	Thr	Ala	Glu	Phe	Lys	Gly	Thr	Phe	Glu	Lys	Ala	Thr	Ser	Glu
			35				40					45			

SUBSTITUTE SHEET

Ala Tyr Ala Tyr Ala Asp Thr Leu Lys Lys Asp Asn Gly Glu Tyr Thr
 50 55 60
 Val Asp Val Ala Asp Lys Gly Tyr Thr Leu Asn Ile Lys Phe Ala Gly
 65 70 75 80
 Lys Glu Lys Thr Pro Glu Glu Pro Lys Glu Glu Val Thr Ile Lys Ala
 85 90 95
 Asn Leu Ile Tyr Ala Asp Gly Lys Thr Gln Thr Ala Glu Phe Lys Gly
 100 105 110
 Thr Phe Glu Glu Ala Thr Ala Glu Ala Tyr Arg Tyr Ala Asp Ala Leu
 115 120 125
 Lys Lys Asp Asn Gly Glu Tyr Thr Val Asp Val Ala Asp Lys Gly Tyr
 130 135 140
 Thr Leu Asn Ile Lys Phe Ala Gly Lys Glu Lys Thr Pro Glu Glu Pro
 145 150 155 160
 Lys Glu Glu Val Thr Ile Lys Ala Asn Leu Ile Tyr Ala Asp Gly Lys
 165 170 175
 Thr Gln Thr Ala Glu Phe Lys Gly Thr Phe Glu Glu Ala Thr Ala Glu
 180 185 190
 Ala Tyr Arg Tyr Ala Asp Leu Leu Ala Lys Glu Asn Gly Lys Tyr Thr
 195 200 205
 Val Asp Val Ala Asp Lys Gly Tyr Thr Leu Asn Ile Lys Phe Ala Gly
 210 215 220
 Lys Glu Lys Thr Pro Glu Glu Pro Lys Glu Glu Val Thr Ile Lys Ala
 225 230 235 240
 Asn Leu Ile Tyr Ala Asp Gly Lys Thr Gln Thr Ala Glu Phe Lys Gly
 245 250 255
 Thr Phe Ala Glu Ala Thr Ala Glu Ala Tyr Arg Tyr Ala Asp Leu Leu
 260 265 270
 Ala Lys Glu Asn Gly Lys Tyr Thr Ala Asp Leu Glu Asp Gly Gly Tyr
 275 280 285
 Thr Ile Asn Ile Arg Phe Ala Gly Lys Lys Val Asp Glu Lys Pro Glu
 290 295 300
 Glu
 305

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 921 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: unknown

SUBSTITUTE SHEET

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Escherichia coli LE392/pHDL, DSM 7054

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

GCGGTAGAAA ATAAAGAAGA AACACCAGAA ACACCAGAAA CTGATTCAGA	50
AGAAGAAGTA ACAATCAAAG CTAACCTAAT CTTTGCAAAT GGAAGCACAC	100
AAACTGCAGA ATTCAAAGGA ACATTTGAAA AAGCAACATC AGAAGCTTAT	150
GCGTATGCAG ATACTTTGAA GAAAGACAAT GGAGAATATA CTGTAGATGT	200
TGCAGATAAA GGTTATACTT TAAATATTAA ATTTGCTGGA AAAGAAAAAA	250
CACCAGAAGA ACCAAAAGAA GAAGTTACTA TTAAAGCAAA CTTAATCTAT	300
GCAGATGGAA AAACACAAAC AGCAGAATTC AAAGGAACAT TTGAAGAAGC	350
AACAGCAGAA GCATACAGAT ATGCAGATGC ATTAAAGAAG GACAATGGAG	400
AATATACAGT AGACGTTGCA GATAAAGGTT ATACTTTAAA TATTAAATTT	450
GCTGGAAAAG AAAAAACACC AGAAGAACCA AAAGAAGAAG TTACTATTAA	500
AGCAAACCTTA ATCTATGCAG ATGGAAAAAC ACAAACAGCA GAATTCAAAG	550
GAACATTTGA AGAAGCAACA GCAGAAGCAT ACAGATATGC TGACTTATTA	600
GCAAAAGAAA ATGGTAAATA TACAGTAGAC GTTGCAGATA AAGGTTATAC	650
TTTAAATATT AAATTTGCTG GAAAAGAAA AACACCAGAA GAACCAAAG	700
AAGAAGTTAC TATTAAAGCA AACTTAATCT ATGCAGATGG AAAA ACTCAA	750
ACAGCAGAGT TCAAAGGAAC ATTTGCAGAA GCAACAGCAG AAGCATACAG	800
ATACGCTGAC TTATTAGCAA AAGAAAATGG TAAATATACA GCAGACTTAG	850
AAGATGGTGG ATACACTATT AATATTAGAT TTGCAGGTAA GAAAGTTGAC	900
GAAAAACCAG AAGAATAATA A	921

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 434 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: unknown

(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

SUBSTITUTE SHEET

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Escherichia coli LE392/pHDLG, DSM 7055

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

Ala Val Glu Asn Lys Glu Glu Thr Pro Glu Thr Pro Glu Thr Asp Ser
 1 5 10 15

Glu Glu Glu Val Thr Ile Lys Ala Asn Leu Ile Phe Ala Asn Gly Ser
 20 25 30

Thr Gln Thr Ala Glu Phe Lys Gly Thr Phe Glu Lys Ala Thr Ser Glu
 35 40 45

Ala Tyr Ala Tyr Ala Asp Thr Leu Lys Lys Asp Asn Gly Glu Tyr Thr
 50 55 60

Val Asp Val Ala Asp Lys Gly Tyr Thr Leu Asn Ile Lys Phe Ala Gly
 65 70 75 80

Lys Glu Lys Thr Pro Glu Glu Pro Lys Glu Glu Val Thr Ile Lys Ala
 85 90 95

Asn Leu Ile Tyr Ala Asp Gly Lys Thr Gln Thr Ala Glu Phe Lys Gly
 100 105 110

Thr Phe Glu Glu Ala Thr Ala Glu Ala Tyr Arg Tyr Ala Asp Ala Leu
 115 120 125

Lys Lys Asp Asn Gly Glu Tyr Thr Val Asp Val Ala Asp Lys Gly Tyr
 130 135 140

Thr Leu Asn Ile Lys Phe Ala Gly Lys Glu Lys Thr Pro Glu Glu Pro
 145 150 155 160

Lys Glu Glu Val Thr Ile Lys Ala Asn Leu Ile Tyr Ala Asp Gly Lys
 165 170 175

Thr Gln Thr Ala Glu Phe Lys Gly Thr Phe Glu Glu Ala Thr Ala Glu
 180 185 190

Ala Tyr Arg Tyr Ala Asp Leu Leu Ala Lys Glu Asn Gly Lys Tyr Thr
 195 200 205

Val Asp Val Ala Asp Lys Gly Tyr Thr Leu Asn Ile Lys Phe Ala Gly
 210 215 220

Lys Glu Lys Thr Pro Glu Glu Pro Lys Glu Glu Val Thr Ile Lys Ala
 225 230 235 240

Asn Leu Ile Tyr Ala Asp Gly Lys Thr Gln Thr Ala Glu Phe Lys Gly
 245 250 255

Thr Phe Ala Glu Ala Thr Ala Glu Ala Tyr Arg Tyr Ala Asp Leu Leu
 260 265 270

SUBSTITUTE SHEET

Ala Lys Glu Asn Gly Lys Tyr Thr Ala Asp Leu Glu Asp Gly Gly Tyr
 275 280 285

Thr Ile Asn Ile Arg Phe Ala Gly Lys Lys Val Asp Glu Lys Pro Glu
 290 295 300

Glu Pro Met Asp Thr Tyr Lys Leu Ile Leu Asn Gly Lys Thr Leu Lys
 305 310 315 320

Gly Glu Thr Thr Thr Glu Ala Val Asp Ala Ala Thr Ala Glu Lys Val
 325 330 335

Phe Lys Gln Tyr Ala Asn Asp Asn Gly Val Asp Gly Glu Trp Thr Tyr
 340 345 350

Asp Asp Ala Thr Lys Thr Phe Thr Val Thr Glu Lys Pro Glu Val Ile
 355 360 365

Asp Ala Ser Glu Leu Thr Pro Ala Val Thr Thr Tyr Lys Leu Val Ile
 370 375 380

Asn Gly Lys Thr Leu Lys Gly Glu Thr Thr Thr Lys Ala Val Asp Ala
 385 390 395 400

Glu Thr Ala Glu Lys Ala Phe Lys Gln Tyr Ala Asn Asp Asn Gly Val
 405 410 415

Asp Gly Val Trp Thr Tyr Asp Asp Ala Thr Lys Thr Phe Thr Val Thr
 420 425 430

Glu Met

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1308 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Escherichis coli L392/pHDLG, DSM 7055

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

GCGGTAGAAA ATAAAGAAGA AACACCAGAA ACACCAGAAA CTGATTCAGA	50
AGAAGAAGTA ACAATCAAAG CTAACCTAAT CTTTGCAAAT GGAAGCACAC	100
AAACTGCAGA ATTCAAAGGA ACATTTGAAA AAGCAACATC AGAAGCTTAT	150
GCGTATGCAG ATACTTTGAA GAAAGACAAT GGAGAATATA CTGTAGATGT	200

SUBSTITUTE SHEET

TGCAGATAAA GGTTATACTT TAAATATTAA ATTTGCTGGA AAAGAAAAAA	250
CACCAGAAGA ACCAAAAGAA GAAGTTACTA TTAAAGCAAA CTTAATCTAT	300
GCAGATGGAA AAACACAAAC AGCAGAATTC AAAGGAACAT TTGAAGAAGC	350
AACAGCAGAA GCATACAGAT ATGCAGATGC ATTAAGAAG GACAATGGAG	400
AATATACAGT AGACGTTGCA GATAAAGGTT ATACTTTAAA TATTAAATTT	450
GCTGGAAAAG AAAAAACACC AGAAGAACCA AAAGAAGAAG TTACTATTAA	500
AGCAAACCTTA ATCTATGCAG ATGGAAAAAC ACAAACAGCA GAATTCAAAG	550
GAACATTTGA AGAAGCAACA GCAGAAGCAT ACAGATATGC TGACTTATTA	600
GCAAAAGAAA ATGGTAAATA TACAGTAGAC GTTGCAGATA AAGGTTATAC	650
TTTAAATATT AAATTTGCTG GAAAAGAAAA AACACCAGAA GAACCAAAG	700
AAGAAGTTAC TATTAAAGCA AACTTAATCT ATGCAGATGG AAAAAGTCAA	750
ACAGCAGAGT TCAAAGGAAC ATTTGCAGAA GCAACAGCAG AAGCATAACG	800
ATACGCTGAC TTATTAGCAA AAGAAAATGG TAAATATACA GCAGACTTAG	850
AAGATGGTGG ATACACTATT AATATTAGAT TTGCAGGTAA GAAAGTTGAC	900
GAAAACCAG AAGAACCCAT GGACACTTAC AAATTAATCC TTAATGGTAA	950
AACATTGAAA GCGGAAACAA CTACTGAAGC TGTTGATGCT GCTACTGCAG	1000
AAAAAGTCTT CAAACAATAC GCTAACGACA ACGGTGTTGA CCGTGAATGG	1050
ACTTACGACG ATGCGACTAA GACCTTTACA GTTACTGAAA AACCAGAAGT	1100
GATCGATGCG TCTGAATTAA CACCAGCCGT GACAACTTAC AAAGTTGTTA	1150
TTAATGGTAA AACATTGAAA GCGGAAACAA CTACTAAAGC AGTAGACGCA	1200
GAAACTGCAG AAAAAGCCTT CAAACAATAC GCTAACGACA ACGGTGTTGA	1250
TGGTGTTTGG ACTTATGATG ATGCGACTAA GACCTTTACG GTAAGTAAA	1300
TGTAATAA	1308

SUBSTITUTE SHEET

Claims

5

10

13

20

23

30

- 2 -

SUBSTITUTE SHEET

— B4
 Lys Glu Lys Thr Pro Glu Glu Pro Lys Glu Glu Val Thr Ile Lys Ala
 225 230 235 240
 5 Asn Leu Ile Tyr Ala Asp Gly Lys Thr Gln Thr Ala Glu Phe Lys Gly
 245 250 255
 Thr Phe Ala Glu Ala Thr Ala Glu Ala Tyr Arg Tyr Ala Asp Leu Leu
 260 265 270
 Ala Lys Glu Asn Gly Lys Tyr Thr Ala Asp Leu Glu Asp Gly Gly Tyr
 275 280 — B5 285
 10 Thr Ile Asn Ile Arg Phe Ala Gly Lys Lys Val Asp Glu Lys Pro Glu
 290 295 300
 Glu

15

and variants, subfragments, multiples or mixtures of the domains B1-B5 having the same binding properties.

2. DNA-sequence, characterized in that it codes for the protein according to Claim 1 and has the following nucleotide sequence:

GCG GTA GAA AAT AAA GAA GAA ACA CCA GAA ACA CCA GAA ACT GAT TCA 28
 GAA GAA GAA GTA ACA ATC AAA GCT AAC CTA ATC TTT GCA AAT GGA AGC 96
 25 ACA CAA ACT GCA GAA TTC AAA GGA ACA TTT GAA AAA GCA ACA TCA GAA 144
 GCT TAT GCG TAT GCA GAT ACT TTG AAG AAA GAC AAT GGA GAA TAT ACT 192
 GTA GAT GTT GCA GAT AAA GGT TAT ACT TTA AAT ATT AAA TTT GCT GGA 240
 30 AAA GAA AAA ACA CCA GAA GAA CCA AAA GAA GAA GTT ACT ATT AAA GCA 288
 AAC TTA ATC TAT GCA GAT GGA AAA ACA CAA ACA GCA GAA TTC AAA GGA 336
 ACA TTT GAA GAA GCA ACA GCA GAA GCA TAC AGA TAT GCA GAT GCA TTA 384
 35 AAG AAG GAC AAT GGA GAA TAT ACA GTA GAC GTT GCA GAT AAA GGT TAT 432
 ACT TTA AAT ATT AAA TTT GCT GGA AAA GAA AAA ACA CCA GAA GAA CCA 480
 AAA GAA GAA GTT ACT ATT AAA GCA AAC TTA ATC TAT GCA GAT GGA AAA 528

SUBSTITUTE SHEET

ACA CAA ACA GCA GAA TTC AAA GGA ACA TTT GAA GAA GCA ACA GCA GAA 576
 GCA TAC AGA TAT GCT GAC TTA TTA GCA AAA GAA AAT GGT AAA TAT ACA 624
 5 GTA GAC GTT GCA GAT AAA GGT TAT ACT TTA AAT ATT AAA TTT GCT GGA 672
 AAA GAA AAA ACA CCA GAA GAA CCA AAA GAA GAA GTT ACT ATT AAA GCA 720
 AAC TTA ATC TAT GCA GAT GGA AAA ACT CAA ACA GCA GAG TTC AAA GCA 768
 10 ACA TTT GCA GAA GCA ACA GCA GAA GCA TAC AGA TAC GCT GAC TTA TTA 816
 GCA AAA GAA AAT GGT AAA TAT ACA GCA GAC TTA GAA GAT GGT GGA TAC 864
 ACT ATT AAT ATT AGA TTT GCA GGT AAG AAA GTT GAC GAA AAA CCA GAA 912
 15 GAA TAATAA 921

3. A hybrid protein, characterized in
 that it includes one or more of the B1-B5-domains ac-
 20 cording to Claim 1 which bind to the light chains in
 immunoglobulins of all classes, and domains which bind
 to heavy chains in immunoglobulin G.

4. A hybrid protein according to Claim 3, character-
 25 ized in that the domains which bind to
 heavy chains in immunoglobulin G are chosen from among
 the C1- and C2-domains in protein G or from among any
 other functionally similar proteins which bind to heavy
 chains in immunoglobulin G, and variants, subfragments,
 30 multiples or mixtures thereof having the same binding
 properties.

5. A hybrid protein according to Claim 4, character-
 35 ized in that the hybrid protein has the
 following amino acid sequence:

SUBSTITUTE SHEET

Ala Val Glu Asn Lys Glu Glu Thr Pro Glu Thr Pro Glu Thr Asp Ser
 1 5 10 15

5 Glu Glu Glu Val Thr Ile Lys Ala Asn Leu Ile Phe Ala Asn Gly Ser
 20 25 30

Thr Gln Thr Ala Glu Phe Lys Gly Thr Phe Glu Lys Ala Thr Ser Glu
 35 40 45

10 Ala Tyr Ala Tyr Ala Asp Thr Leu Lys Lys Asp Asn Gly Glu Tyr Thr
 50 55 60

Val Asp Val Ala Asp Lys Gly Tyr Thr Leu Asn Ile Lys Phe Ala Gly
 65 70 75 80

Lys Glu Lys Thr Pro Glu Glu Pro Lys Glu Glu Val Thr Ile Lys Ala
 85 90 95

15 Asn Leu Ile Tyr Ala Asp Gly Lys Thr Gln Thr Ala Glu Phe Lys Gly
 100 105 110

Thr Phe Glu Glu Ala Thr Ala Glu Ala Tyr Arg Tyr Ala Asp Ala Leu
 115 120 125

20 Lys Lys Asp Asn Gly Glu Tyr Thr Val Asp Val Ala Asp Lys Gly Tyr
 130 135 140

Thr Leu Asn Ile Lys Phe Ala Gly Lys Glu Lys Thr Pro Glu Glu Pro
 145 150 155 160

25 Lys Glu Glu Val Thr Ile Lys Ala Asn Leu Ile Tyr Ala Asp Gly Lys
 165 170 175

Thr Gln Thr Ala Glu Phe Lys Gly Thr Phe Glu Glu Ala Thr Ala Glu
 180 185 190

Ala Tyr Arg Tyr Ala Asp Leu Leu Ala Lys Glu Asn Gly Lys Tyr Thr
 195 200 205

30 Val Asp Val Ala Asp Lys Gly Tyr Thr Leu Asn Ile Lys Phe Ala Gly
 210 215 220

Lys Glu Lys Thr Pro Glu Glu Pro Lys Glu Glu Val Thr Ile Lys Ala
 225 230 235 240

35 Asn Leu Ile Tyr Ala Asp Gly Lys Thr Gln Thr Ala Glu Phe Lys Gly
 245 250 255

SUBSTITUTE SHEET

Thr Phe Ala Glu Ala Thr Ala Glu Ala Tyr Arg Tyr Ala Asp Leu Leu
 260 265 270
 5 Ala Lys Glu Asn Gly Lys Tyr Thr Ala Asp Leu Glu Asp Gly Gly Tyr
 275 280 285
 Thr Ile Asn Ile Arg Phe Ala Gly Lys Lys Val Asp Glu Lys Pro Glu
 290 295 300
 10 ~~Glu Pro Met~~ Asp Thr Tyr Lys Leu Ile Leu Asn Gly Lys Thr Leu Lys
 305 310 315 320
 Gly Glu Thr Thr Thr Glu Ala Val Asp Ala Ala Thr Ala Glu Lys Val
 325 330 335
 15 Phe Lys Gln Tyr Ala Asn Asp Asn Gly Val Asp Gly Glu Trp Thr Tyr
 340 345 350
 Asp Asp Ala Thr Lys Thr Phe Thr Val Thr Glu Lys Pro Glu Val Ile
 355 360 365
 Asp Ala Ser Glu Leu Thr Pro Ala Val Thr Thr Tyr Lys Leu Val Ile
 370 375 380
 20 Asn Gly Lys Thr Leu Lys Gly Glu Thr Thr Thr Lys Ala Val Asp Ala
 385 390 395 400
 Glu Thr Ala Glu Lys Ala Phe Lys Gln Tyr Ala Asn Asp Asn Gly Val
 405 410 415
 25 Asp Gly Val Trp Thr Tyr Asp Asp Ala Thr Lys Thr Phe Thr Val Thr
 420 425 430
 Glu Met

30 and variants, subfragments, multiples or mixtures of the
 domains B1-B5 having the same binding properties.

35 6. DNA-sequence, characterized in that
 it codes for a protein according to Claim 5 and has the
 following nucleotide sequence:

GCG GTA GAA AAT AAA GAA GAA ACA CCA GAA ACA CCA GAA ACT GAT TCA 48
GAA GAA GAA GTA ACA ATC AAA GCT AAC CTA ATC TTT GCA AAT GGA AGC 96
ACA CAA ACT GCA GAA TTC AAA GGA ACA TTT GAA AAA GCA ACA TCA GAA 144
GCT TAT GCG TAT GCA GAT ACT TTG AAG AAA GAC AAT GGA GAA TAT ACT 192
5 GTA GAT GTT GCA GAT AAA GGT TAT ACT TTA AAT ATT AAA TTT GCT GGA 240
AAA GAA AAA ACA CCA GAA GAA CCA AAA GAA GAA GTT ACT ATT AAA GCA 288
AAC TTA ATC TAT GCA GAT GGA AAA ACA CAA ACA GCA GAA TTC AAA GGA 336
10 ACA TTT GAA GAA GCA ACA GCA GAA GCA TAC AGA TAT GCA GAT GCA TTA 384
AAG AAG GAC AAT GGA GAA TAT ACA GTA GAC GTT GCA GAT AAA GGT TAT 432
ACT TTA AAT ATT AAA TTT GCT GGA AAA GAA AAA ACA CCA GAA GAA CCA 480
AAA GAA GAA GTT ACT ATT AAA GCA AAC TTA ATC TAT GCA GAT GGA AAA 528
15 ACA CAA ACA GCA GAA TTC AAA GGA ACA TTT GAA GAA GCA ACA GCA GAA 576
GCA TAC AGA TAT GCT GAC TTA TTA GCA AAA GAA AAT GGT AAA TAT ACA 624
GTA GAC GTT GCA GAT AAA GGT TAT ACT TTA AAT ATT AAA TTT GCT GGA 672
20 AAA GAA AAA ACA CCA GAA GAA CCA AAA GAA GAA GTT ACT ATT AAA GCA 720
AAC TTA ATC TAT GCA GAT GGA AAA ACT CAA ACA GCA GAG TTC AAA GGA 768
ACA TTT GCA GAA GCA ACA GCA GAA GCA TAC AGA TAC GCT GAC TTA TTA 816
GCA AAA GAA AAT GGT AAA TAT ACA GCA GAC TTA GAA GAT GGT GGA TAC 864
25 ACT ATT AAT ATT AGA TTT GCA GGT AAG AAA GTT GAC GAA AAA CCA GAA 912
GAA CCC ATG GAC ACT TAC AAA TTA ATC CTT AAT GGT AAA ACA TTG AAA 960
GGC GAA ACA ACT ACT GAA GCT GTT GAT GCT GCT ACT GCA GAA AAA GTC 1008
30 TTC AAA CAA TAC GCT AAC GAC AAC GGT GTT GAC GGT GAA TGG ACT TAC 1056
GAC GAT GCG ACT AAG ACC TTT ACA GTT ACT GAA AAA CCA GAA GTG ATC 1104
GAT GCG TCT GAA TTA ACA CCA GCC GTG ACA ACT TAC AAA CTT GTT ATT 1152
AAT GGT AAA ACA TTG AAA GGC GAA ACA ACT ACT AAA GCA GTA GAC GCA 1200
35 GAA ACT GCA GAA AAA GCC TTC AAA CAA TAC GCT AAC GAC AAC GGT GTT 1248
GAT GGT GTT TGG ACT TAT GAT GAT GCG ACT AAG ACC TTT ACG GTA ACT 1296
GAA ATG TAATAA 1308

SUBSTITUTE SHEET

7. DNA-sequence, characterized in that it codes for a protein according to Claim 3 and 4.

5 8. DNA-sequence, characterized in that it hybridizes to the DNA-sequence of Claim 2, 6 or 7 under conventional conditions and codes for a protein which has the same binding properties as the protein according to any one of Claims 1 and 3-5.

10 9. A plasmid vector, characterized in that it includes a DNA-sequence according to any one of Claims 2 and 6-8, preferably the vector pHDLG or pHDL.

15 10. A host cell, characterized in that it is transformed with the hybrid plasmid according to Claim 9, in particular a host which belongs to the species E. coli, particularly E. coli LE392, or Bacillus subtilis, Saccaromyces cerevisiae, preferably Id. Ref. DSSM E. coli LE392 pHDL and E. coli LE392/pHDLG respectively.

20

11. A method for producing a protein according to Claims 1 and 3-5, characterized by cultivating a host cell according to Claim 10 under suitable conditions; accumulating the protein in the culture or
25 lysing the cells and extracting the protein therefrom.

12. A reagent apparatus for binding, separating and identifying immunoglobulins, characterized
30 in that it includes a protein according to any one of Claims 1 and 3-5.

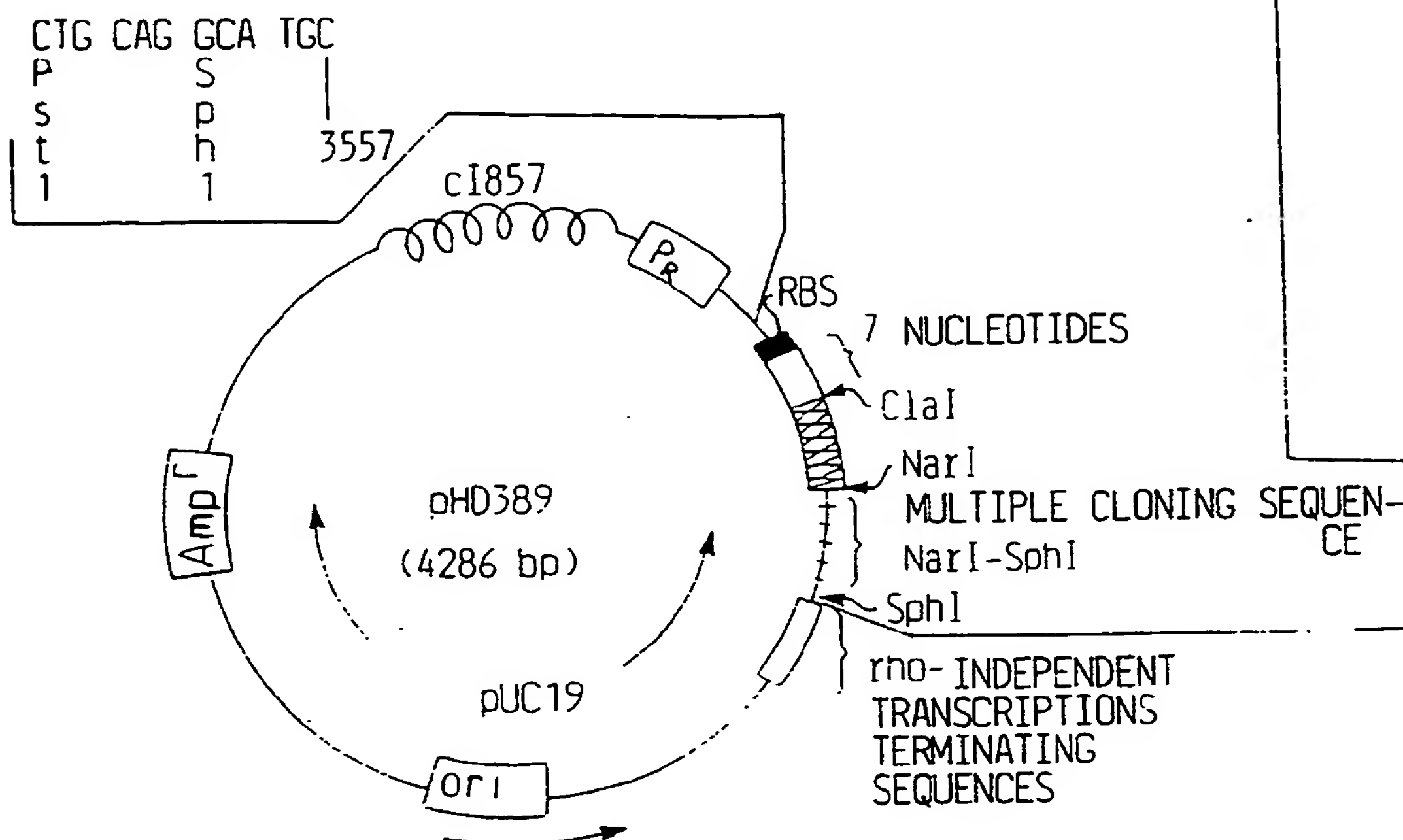
13. A composition, characterized in that it includes a protein according to any one of Claims 1
35 and 3-5, and optionally additives or carriers.

SUBSTITUTE SHEET

14. A pharmaceutical composition, c h a r a c -
t e r i z e d in that it includes a protein according
to any one of Claims 1 and 3-5, and optionally a pharma-
ceutically acceptable carrier or extender.

SUBSTITUTE SHEET

3406
 1 / 18
 AAGCTTAAGGAGGTTAATCG ATG AAA AAA ACT GCT ATC GCT ATC GCT GTT
 H A C met
 i f RBS l
 n 1 a
 3 2 1
 GCT CTG GCT GGT TTC GCT ACT GTT GCT CAG GCG GCG CCG AGA TCT
 aia N B
 a r 1 2
 AAA CAG GAA TTC GAG CTC GGT ACC CGG GGA TCC TCT AGA GTC GAC
 E S K X B X S
 C a c p m a b b a l
 O I 1 1 1 1 1 1



■■■■■ SIGNAL PEPTIDE FOR THE SEQUENCE FROM ompA

■ RBS = RIBOSOMAL BINDING SEQUENCE AGGAGG

P_R = "RIGHT" PROMOTOR FROM COLIPHAGE λ

cI857 THE GENE FOR A HEAT-SENSITIVE REPRESSOR-PROTEIN FROM COLIPHAGE λ

FIG. 1 PLASMA pHD 389. THE RIBOSOMAL BINDING-SEQUENCE (EMPHASIZED WITH A FULL LINE), THE SEQUENCE FOR SIGNAL PEPTIDE FROM ompA (FROM E.coli) (DOTTED LINE) AND RECOGNITION SEQUENCE FOR SEVERAL RESTRICTION ENZYMES ARE SHOWN.

MULTIPLE CLONING SEQUENCE

NarI
 BglII
 EcoRI
 SacI
 KpnI
 XmaI
 BamHI
 XbaI
 SalI
 PstI
 SphI

UNIQUE CLEAVING SITES

SUBSTITUTE SHEET

[illegible]

FIG. 2
(CONT.)

3/18

GCATACAGATATGCAGATGCATTAAAGAGCAATGGAGAAATATACAGTAGACGTTGCA 420
AlaTyrArgTyrAlaAspAlaLeuLysLysAspAsnGlyGluTyrThrValAspValAla 140
GATAAAGGTTATACTTAAATATTAATTTGCTGGAATAAGAAAAACACCAGAAACCA 480
AspLysGlyTyrThrLeuAsnIleLysPheAlaGlyLysGluLysThrProGluPro 160
AAAGAAGAAGTTACTATTAAGCAAACCTTAATCTATGCAGATGGAAAAACACAAACAGCA 540
LysGluGluValThrIleLysAlaAsnLeuIleTyrAlaAspGlyLysThrGlnThrAla 180
GAATTCAAAGGAACATTTGAAGAGCAACAGCAGACATACAGATATGCTGACTTATTA 600
GluPheLysGlyThrPheGluGluAlaThrAlaGluAlaTyrArgTyrAlaAspLeuLeu 200
GCAAAAGAAAATGGTAAATATACAGTAGACGTTGCAGATAAAGGTTATACTTTAAATATT 660
AlaLysGluAsnGlyLysTyrThrValaspValAlaAspLysGlyTyrThrLeuAsnIle 220
AAATTGCTGGAATAAGAAAAACACCAGAAGAACCAAGAAAGTACTATAAAGCA 720
LysPheAlaGlyLysGluLysThrProGluGluProLysGluGluValThrIleLysAla 240
AACTTAATCTATGCAGATGGAAAAACTCAAACAGCAGAGTTCAAAGGAACATTTGCAGAA 780
AsnLeuIleTyrAlaAspGlyLysThrGlnThrAlaGluPheLysGlyThrPheAlaGlu 260

FIG. 2
(CONT.)

SUBSTITUTE SHEET

4 / 18

FIG. 2
(CONT.)

SUBSTITUTE SHEET

V	V	V	V	V	V	1200
AAACTTGTATTAATGGTAAACATTGAAAGCGGAAACAACACTAAAGCAGTAGACGCA						400
LysLeuValIleAsnGlyLysThrLeuLysGlyGluThrThrLysAlaValAspAla						
V	V	V	V	V	V	1260
GAAACTGCAGAAAGCCCTTCAAACAATACGCTAACGACAACGGTGTGATGGTGTGG						420
GluThrAlaGluLysAlaPheLysGlnTyrAlaAsnAspAsnGlyValAspGlyValTrp						
V	V	V	V	V	V	1308
ACTTATGATGATGCGACTAAGACCTTTACGGTAACTGAAATGTAAATAA						434
ThrTyrAspAspAlaThrLysThrPheThrValThrGluMet - -						

FIG. 2

6/18

FIG.3 SCHEMATIC OVERALL VIEW OF THE PRODUCTION OF PROTEIN L

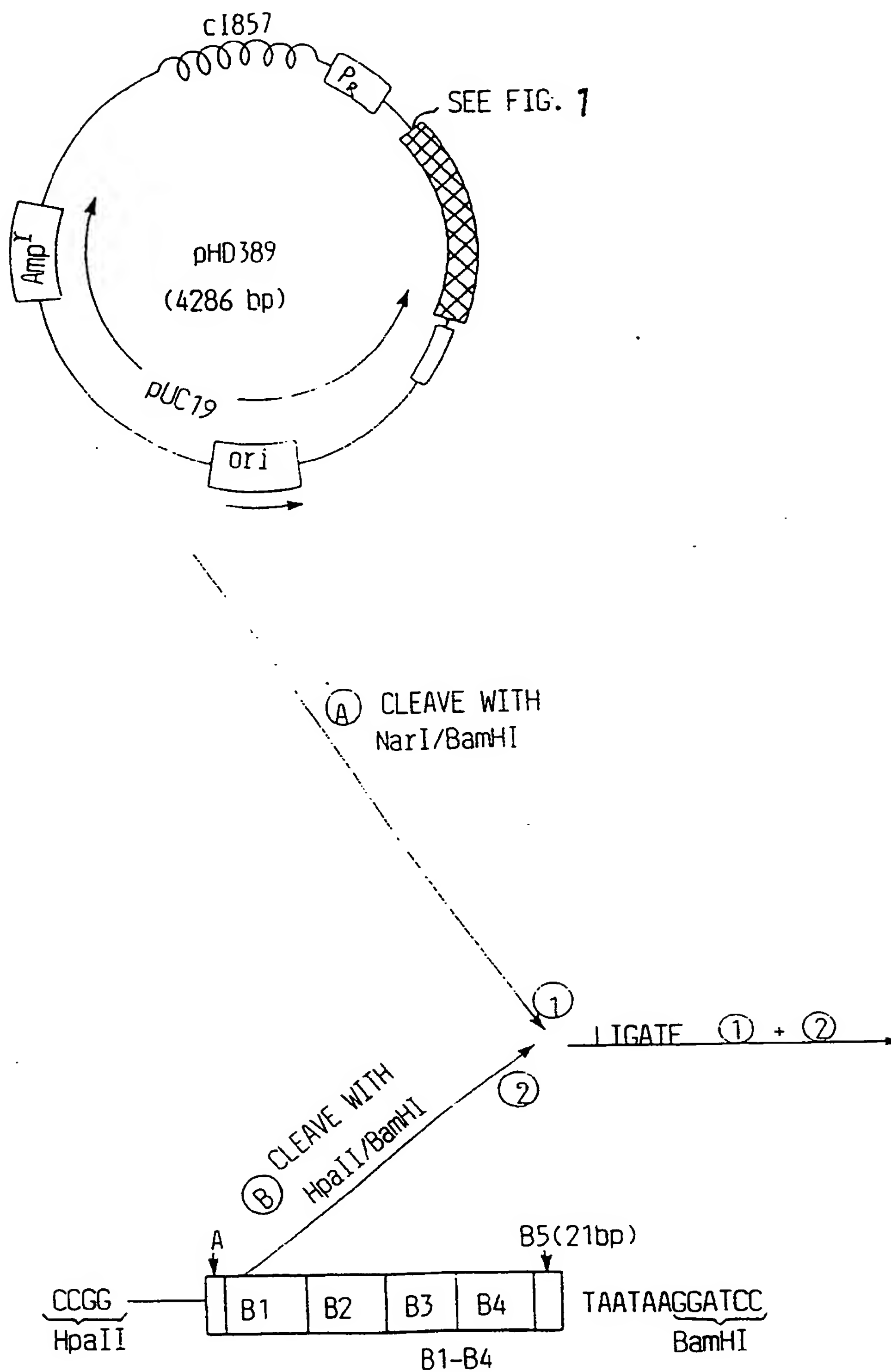
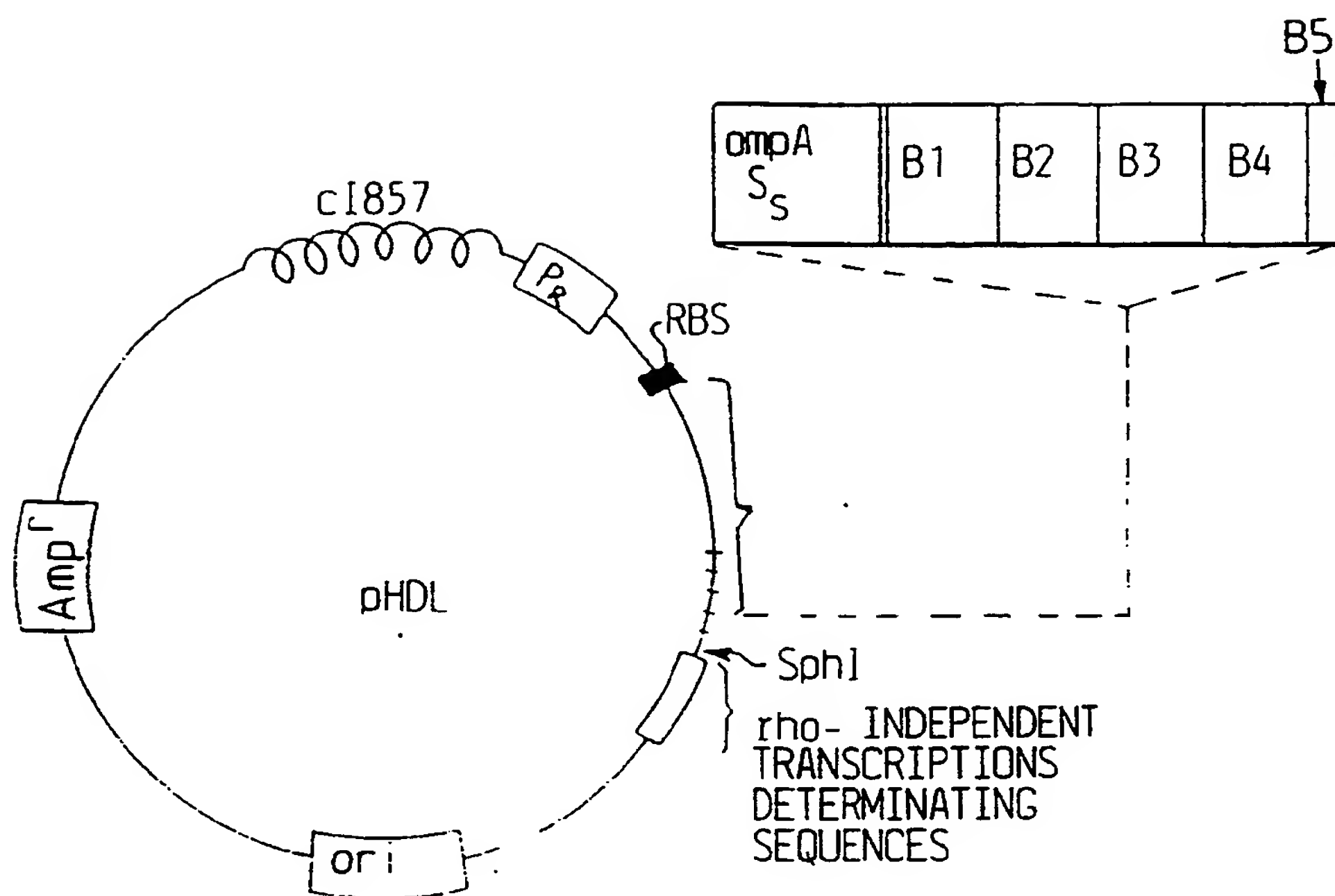


FIG.3(1)

SUBSTITUTE SHEET

7/18



TRANSFORM
TO E.coli
STRAIN LE392

FIG.3(2)

SUBSTITUTE SHEET

8/18

FIG. 4

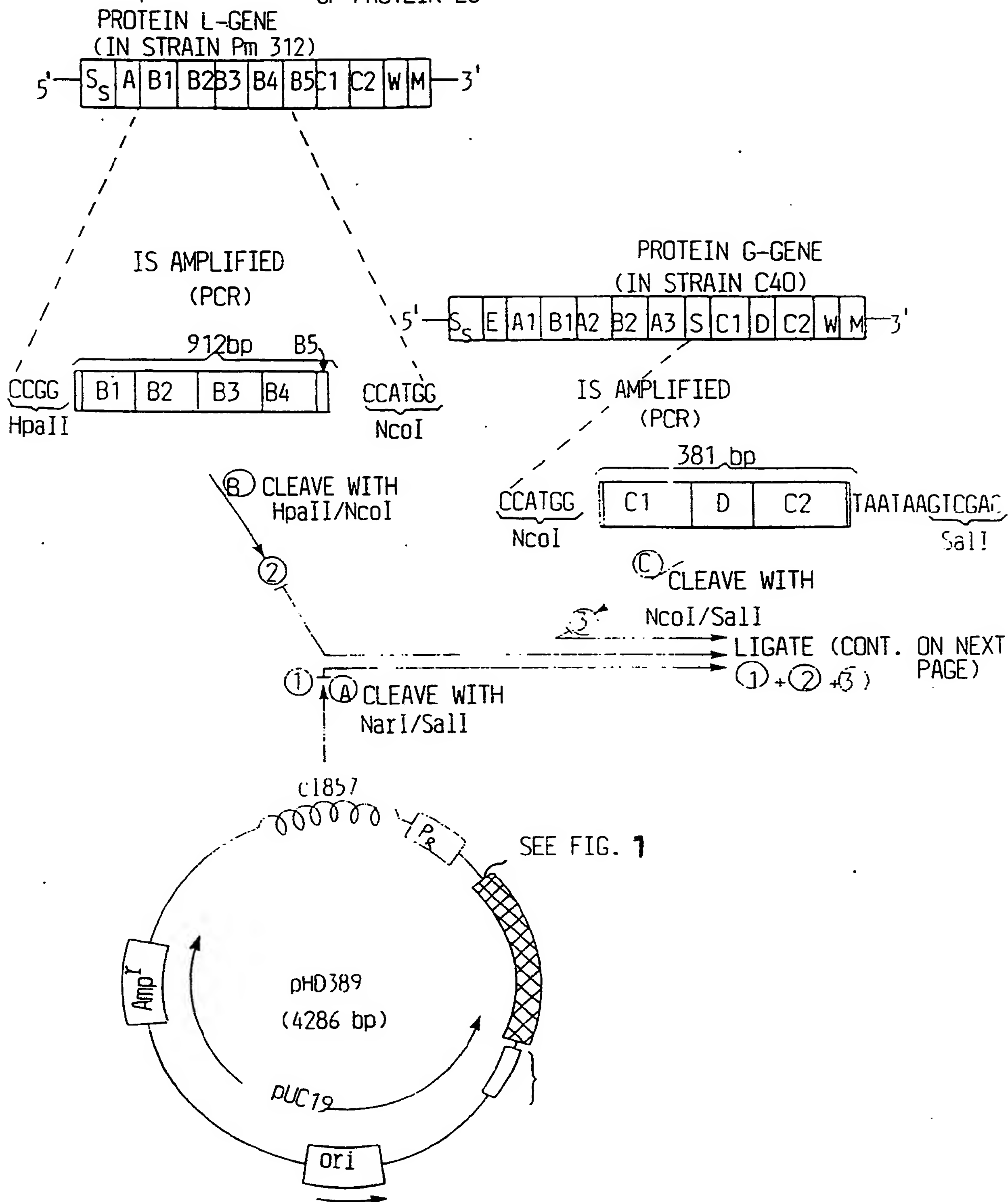
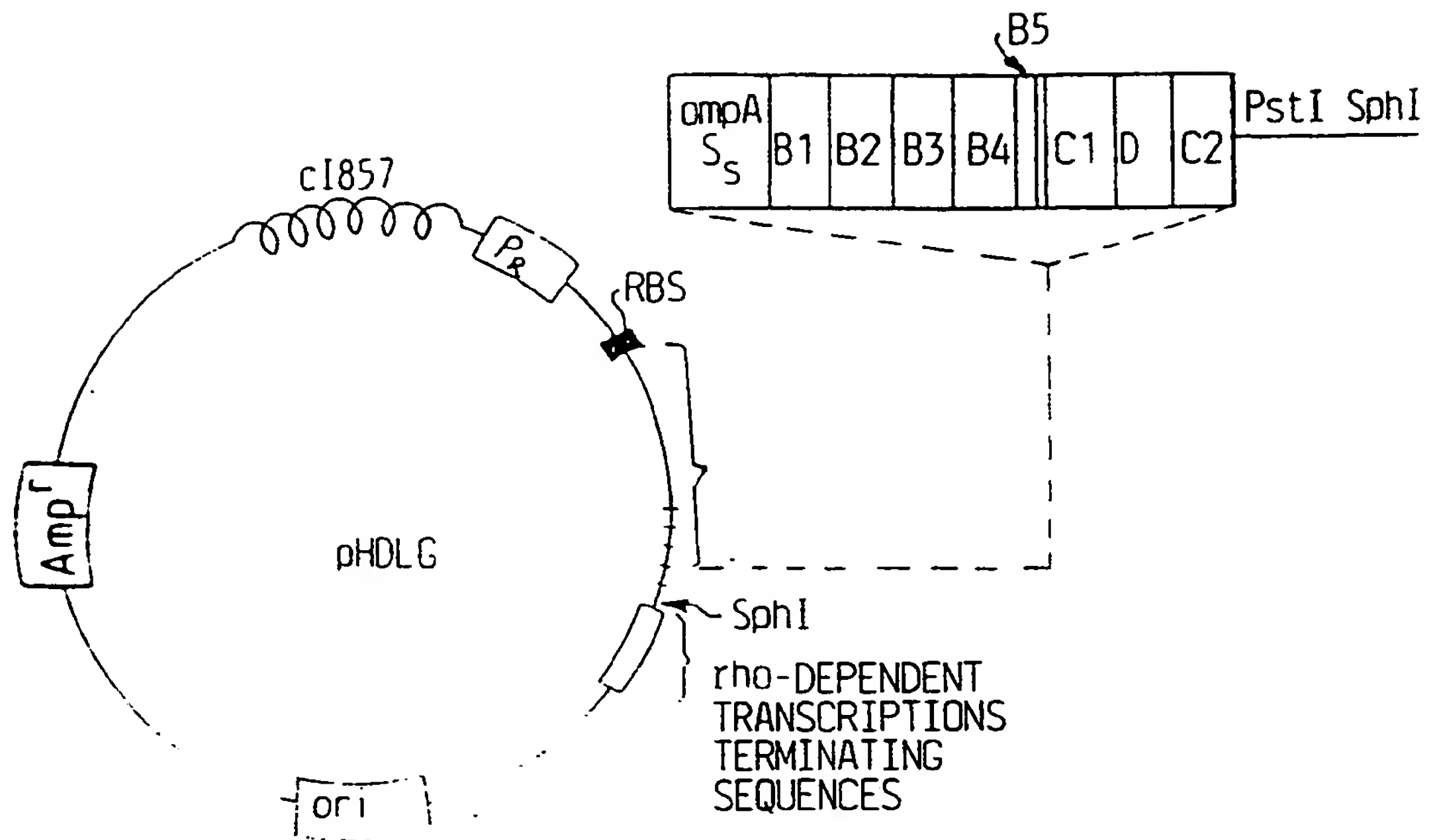
SCHEMATIC OVERALL VIEW OF PRODUCTION
OF PROTEIN LG

FIG.4(1)

SUBSTITUTE SHEET

9/18



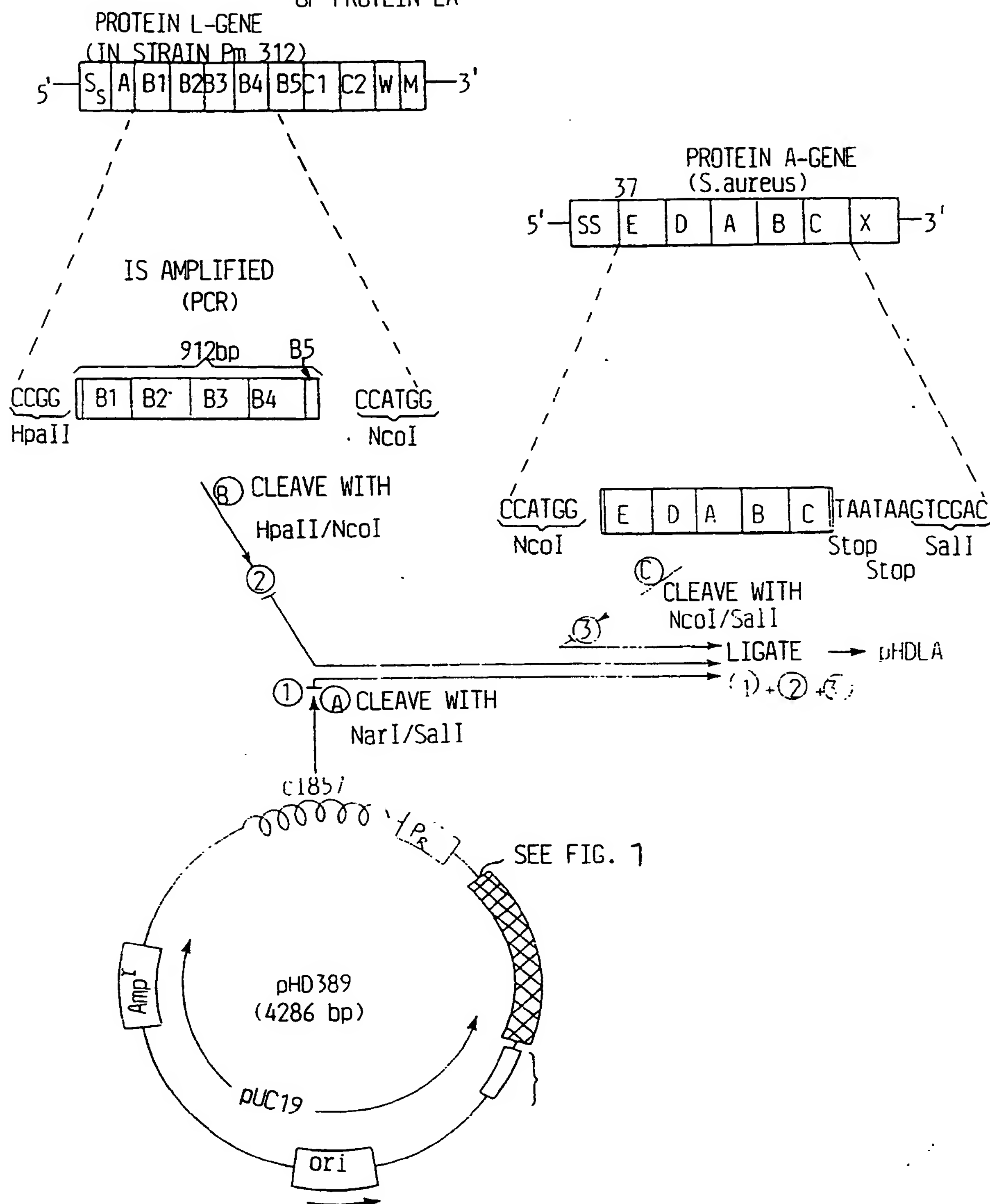
↓
TRANSFORM
TO *E.coli*
STRAIN LE392

FIG.4(2)

SUBSTITUTE SHEET

.10/18

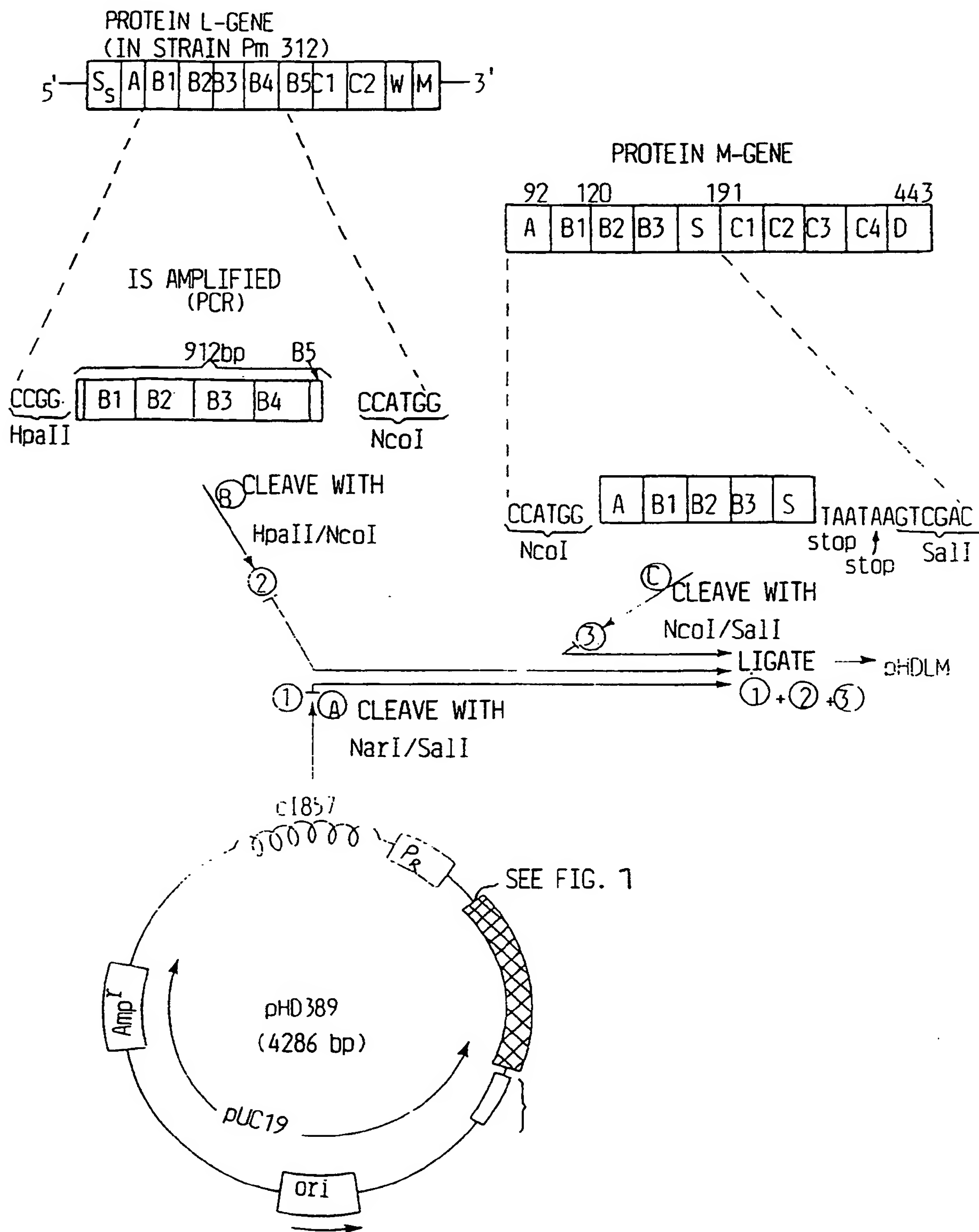
FIG. 5a SCHEMATIC OVERALL VIEW OF THE PRODUCTION OF PROTEIN LA



SUBSTITUTE SHEET

11/18

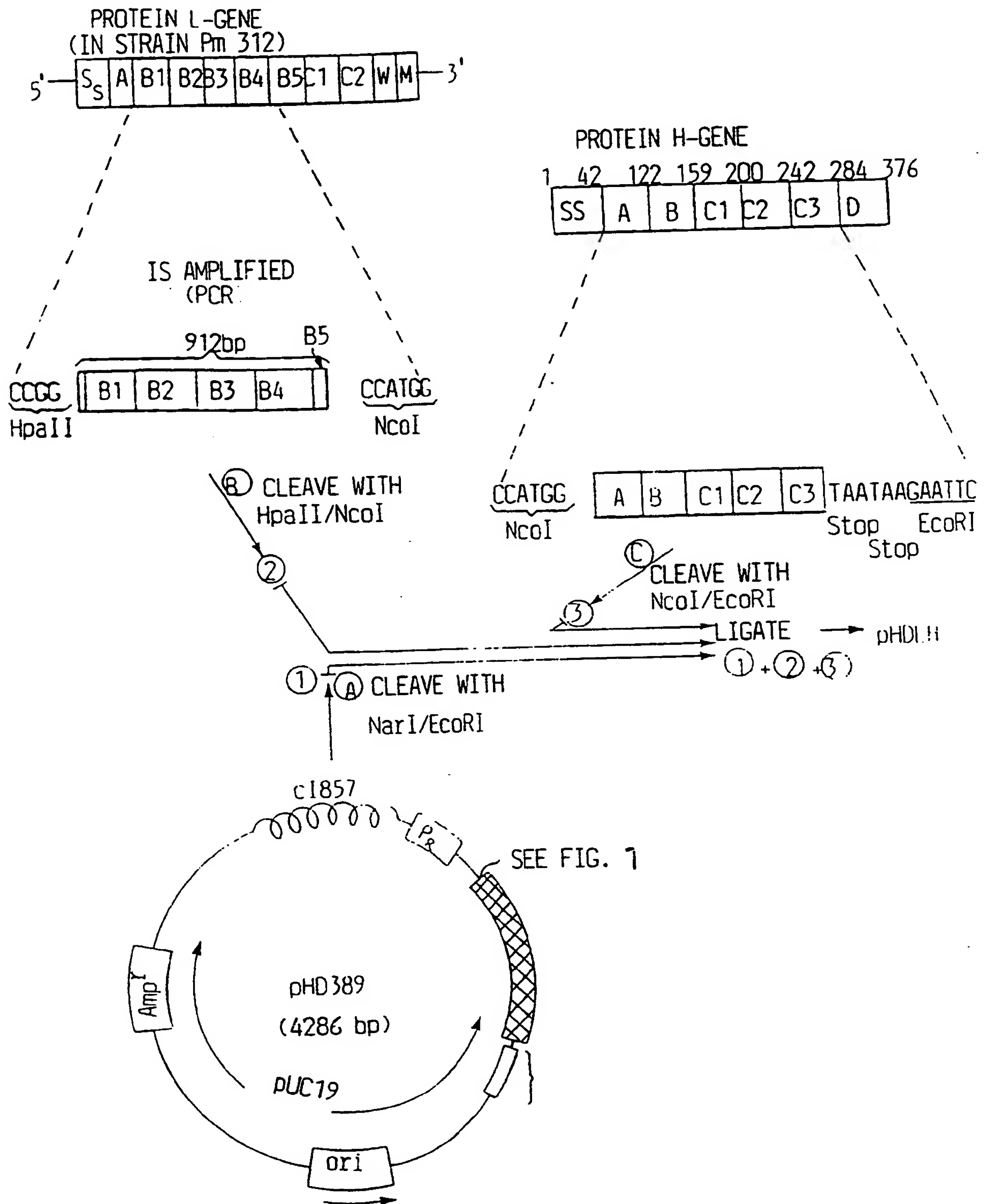
FIG. 5b SCHEMATIC OVERALL VIEW OF THE PRODUCTION OF PROTEIN LM



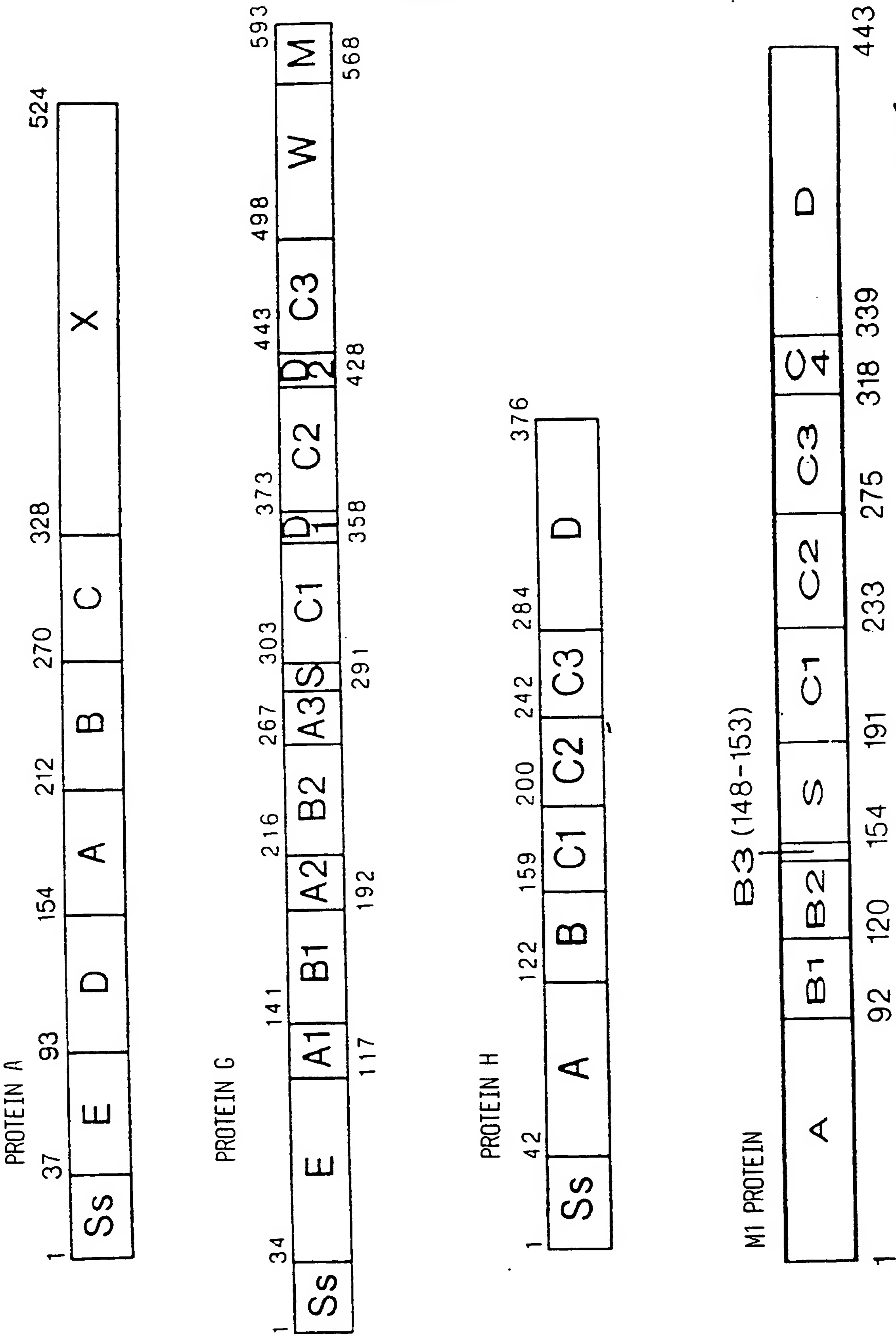
SUBSTITUTE SHEET

12/18

FIG. 5c SCHEMATIC OVERALL VIEW OF THE PRODUCTION OF PROTEIN LH



SUBSTITUTE SHEET



<p> ^V AACGGTGATGGTAA^VTCCTAGGGAAGTTATAGAGATCTT^VGCAGCAACAATCCCGCAATA ^V AsnGlyAspGlyAsnProArgGluValIleGluAspLeuAlaIleAsnAsnProAlaIle </p>	60 20
<p> ^V CAAAATATACGTTTACGTCACGAAACAAGGACTTAAAGCGGAGATTAGAGAAATGCAATG ^V GlnAsnIleArgLeuArgHisGluAsnLysAspLeuLysAlaArgLeuGluAsnAlaMet </p>	120 40
<p> ^V GAAGTTGCAGGAGAGATTTTAAAGAGAGCTGAAGAACTTGAAAGCAAAACAAGCCTTA ^V GluValAlaGlyArgAspPheLysArgAlaGluGluLeuGluLysAlaLysGlnAlaLeu </p>	180 60
<p> ^V GAAGACCAGCGTAAGATTTAGAACTAAATTAAAGAACTACAACAAGACTATGACTTA ^V GluAspGlnArgLysAspLeuGluThrLysLeuLysGluLeuGlnGlnAspTyrAspLeu </p>	240 80
<p> ^V GCAAAGGAATCAACAAGTTGGGATAGACAAAGACTTGAA^VAAAGAGTTAGAGAGAAAG ^V AlaLysGluSerThrSerTrpAspArgGlnArgLeuGluLysGluLeuGluGluLysLys </p>	300 100
<p> ^V GAAGCTCTTGAAATTAGCGATAGACCAGGCAAGTCGGGACTACCATAGAGCTACCGCTTTA ^V GluAlaLeuGluLeuAlaIleAspGlnAlaSerArgAspTyrHisArgAlaThrAlaLeu </p>	360 120
<p> ^V GAAAAAGAGTTAGAGAGAA^VAAAGAAAGCTCTTGAAATTAGCGATAGACCAGGAGTCAG ^V GluLysGluLeuGluGluLysLysAlaLeuGluLeuAlaIleAspGlnAlaSerGln </p>	420 140
<p> ^V GACTATAATAGAGCTAACGTCTTAGAAAAAGAGTTAGAAACGATTACTAGAGACAAGAG ^V AspTyrAsnArgAlaAsnValLeuGluLysGluLeuGluThrIleThrArgGluGlnGlu </p>	480 160
<p> ^V ATTAATCGTAATCTTTTAGGCAATGCAA^VAAACTTGAACTTGATCAACTTTCATCTGAAAAA ^V IleAsnArgAsnLeuLeuGlyAsnAlaLysLeuGluLeuAspGlnLeuSerSerGluLys </p>	540 180

FIG. 7 (CONT.)

FIG. 7 (CONT.)

Amino acid sequence and nucleic acid sequence for protein M1, IgG-binding somewhere between amino acid 1-190.

840
290
900
300
960
320
1020
340
1080
360
1140
380
1200
400
1260
420
1320
440

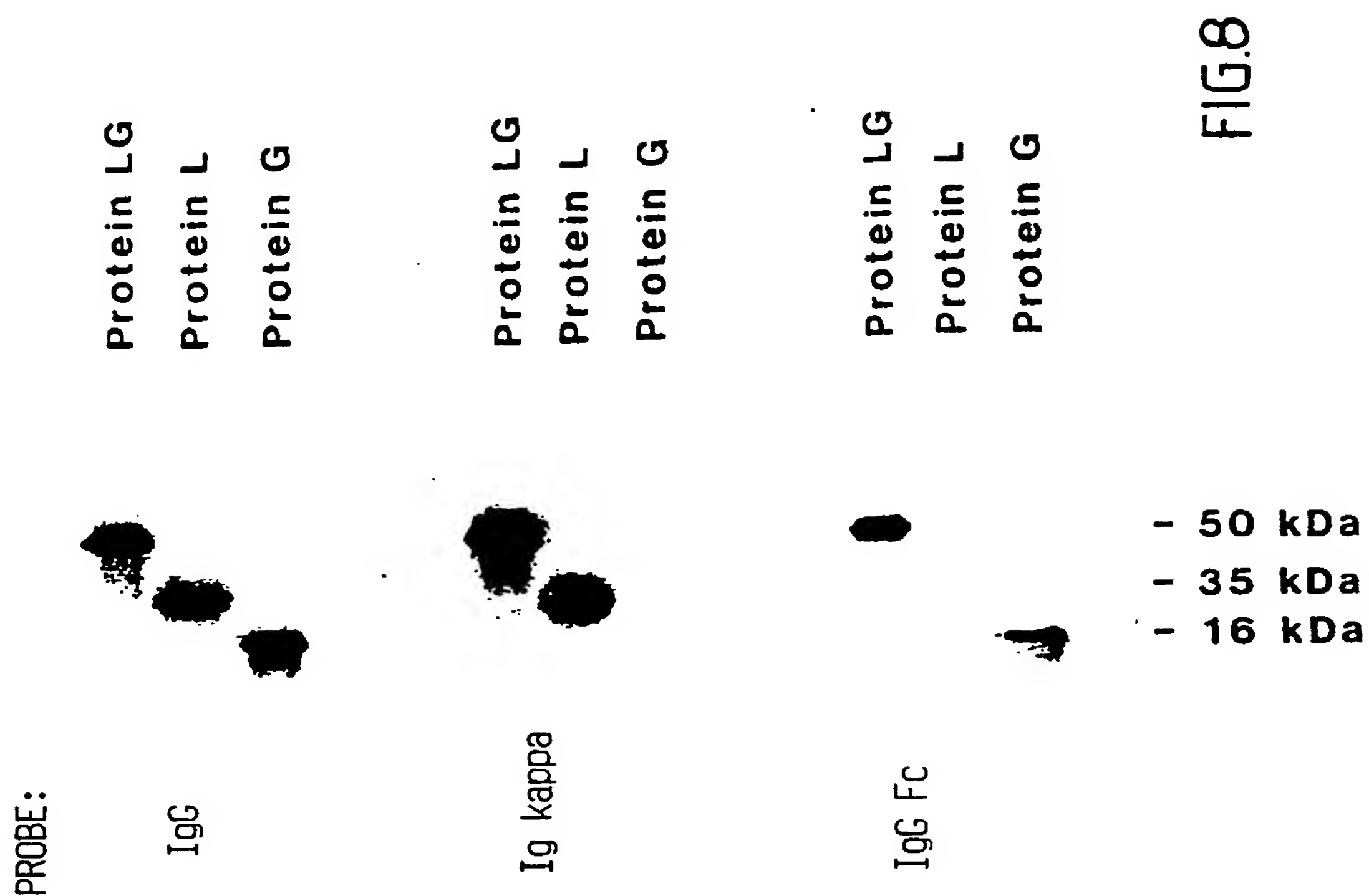
GAAAAGATTAGCAAACTTGACTGCTGAACCTTGATAAGGTTAAAGAGAGAAAACAATC
GluLysAspLeuAlaAsnLeuThrAlaGluLeuAspLysValLysGluLysGlnIle
TCAGACGCAAGCCGTCAACGGCTTCGCCGTGACTTGACGTCATCACGTGAGCTAAGAAA
SerAspAlaSerArgGlnArgLeuArgAspLeuAspAlaSerArgGluAlaLysLys
CAAGTTGAAAAAGCTTTAGAAGAGCAACAGCAAAATTAGCTGCTCTTGAAAAACTTAAC
GlnValGluLysAlaLeuGluGluAlaAsnSerLysLeuAlaAlaLeuGluLysLeuAsn
AAAGAGCTTGAGAGAGCAAGAAATTAAACAGAAAGAAAGCTGAACCTACAAGCAAAA
LysGluLeuGluGluSerLysLysLeuThrGluLysGluLysAlaGluLeuGlnAlaLys
CTTGAGCAGAGCAAAAGCACTCAAGAGAACAAATTAGCGAAACAAGCTGAAGAACTCGCA
LeuGluAlaGluAlaLysAlaLeuLysGluGlnLeuAlaLysGlnAlaGluLeuAla
AACTAAGAGCTGGAAAAGCATCAGACTCACAAACCCCTGATACAAACCAGGAACAAA
LysLeuArgAlaGluLysAlaSerAspSerGlnThrProAspThrLysProGluAsnLys
GCTGTTCCAGGTAAAGGTCAAGCACCAAGCAGGTACAAACCTAACCAAAACAAGCA
AlaValProGluLysGluGlnAlaProGlnAlaGluThrLysProAsnGlnAsnLysAla
CCAATGAAGGAACTAAGAGACAGTTACCATCAACAGGTGAACAGCTAACCCATTCTTC
ProMetLysGluThrLysArgGlnLeuProSerThrGluThrAlaAsnProPhe
ACAGCGCACGCGTTACTGTTATGGCAACAGCTGGAGTAGCAGCAGTTGTAAACGCAAA
ThrAlaAlaArgValThrValMetAlaThrAlaGluValAlaValLysArgLys

FIG.7

GAAGAAAACCTAA 1329
GluGluAsn>>> 443

SUBSTITUTE SHEET

17/18



SUBSTITUTE SHEET

BEST AVAILABLE COPY

18/18

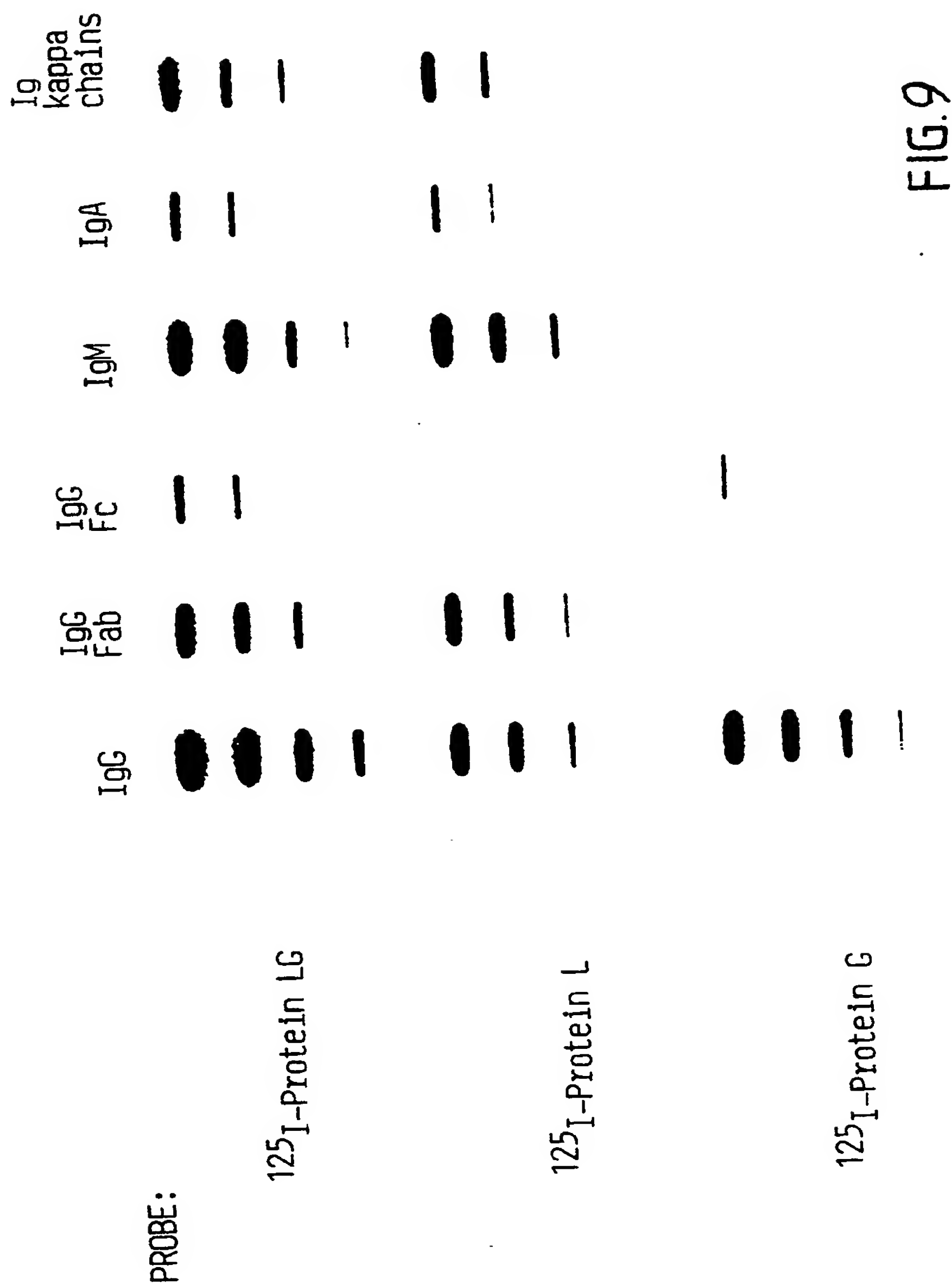


FIG. 9

SUBSTITUTE SHEET

BEST AVAILABLE COPY

A. CLASSIFICATION OF SUBJECT MATTER

IPC5: C07K 13/00, C12N 15/31, C12N 15/62, A61K 37/02, C07K 3/18
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC5: C07K, C12N, A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

SE,DK,FI;NO classes as above

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	EP, A2, 0255497 (HIGHECH RECEPTOR AB), 3 February 1988 (03.02.88) --	1-2,8-14
Y	WO, A1, 8705631 (PHARMACIA AB), 24 Sept 1987 (24.09.87), see especially claim 9 --	3-14
P,X	The Journal of Biological Chemistry, Volume 267, No 18, 1992, William Kastern et al, "Structure of Peptostreptococcal Protein L and Identification of a Repeated Immunoglobulin Light Chain-binding Domain", pp. 12820-12825 --	1-2,8-14

☒ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents:

- "A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier document but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed

- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance: the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search

21 July 1993

Date of mailing of the international search report

29 -07- 1993

Name and mailing address of the ISA/
Swedish Patent Office
Box 5055, S-102 42 STOCKHOLM
Facsimile No. +46 8 666 02 86

Authorized officer

Mikael G:son Bergstrand
Telephone No. +46 8 782 25 00

2

INTERNATIONAL SEARCH REPORT

International application No.
 PCT/SE 93/00375

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	INFECTION AND IMMUNITY, 58(1990-05):5 William Kastern et al: "Protein L, a Bacterial Immunoglobulin-Binding Protein and Possible Virulence Determinant", page 1217 - page 1222; see especially fig. 4 and 5	1-2,8-14
Y	-----	3-14

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT
Information on patent family members

02/07/93

International application No.
PCT/SE 93/00375

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP-A2- 0255497	03/02/88	JP-A- 63032372 US-A- 4876194	12/02/88 24/10/89
WO-A1- 8705631	24/09/87	DE-A- 3783191 EP-A,B- 0262192 SE-T3- 0262192	04/02/93 06/04/88

Form PCT/ISA/210 (patent family annex) (July 1992)

THIS PAGE BLANK (USPT